Authors' responses to reviewer's comments follow. A copy of the reviewer comment is given (with comment 'number') followed by a response (blue font).

Response to referee 2

1. General comments

   This is a well-written paper with interesting data concerning atmospheric observations and validations of fossil CO2, and CO and SF6 emissions from Korea, and the Asian main land. Upon reading, I have made notes and comments that I present below. A more general remark (also given below) is that I invite/encourage the authors to more explicitly conclude what their observations show concerning the quality of the inventories, and if these inventories are thrus worthy or not. By drawing conclusions in that style, these data will be more accessible and valuable to policy makers, and might help to improve the inventories. I recommend publication after the authors have dealt with my comments below.

   We thank you for your comments on this paper's value. We also appreciate your helpful comments to improve our manuscript. According to your specific comments, we revised our manuscript.

2. page 5, lines 108-109. at what flow rate are the flasks filled, or rather: is the flask air composition an average over some period of time, or merely a point in time?.

   Samples are collected with a semi-automated sampler that first flushes the flasks at 5-6 L/min for 10 minutes then pressurizes them to 5.5 psig (5-6 LPM) in less than 1 min. Therefore, samples are integrated over 1-2 min.

   We added the sentence on line 114.

   Line 114: Two pairs of flask-air samples (4 flasks total, 2 L, borosilicate glass with Teflon O-ring sealed stopcocks) were collected about weekly from a 40 m tall tower at AMY, regardless of wind direction and speed from May 2014 to August 2016, generally between 1400 to 1600 local time (Table

S1) using a semi-automated portable sampler. A pair of flasks was flushed for 10 min at 5-6 L min^{-1} then pressurized to 5.5 psig in less than 1 min. A second pair is collected shortly after the first (within 20 min). The portable sampler was checked for leaks after pressurizing by observing the pressure gauge before closing the stopcocks, Batches of sampled flasks were shipped to Boulder, CO, USA every two months

3. line 107 "Two pairs of flask-air samples (4 flasks total," In tabel S1 I see only one value per week. Is this an average? Flasks taken together for 14C mm?

We collect 4 samples but 2 among them were analyzed for ^{14}C in CO_2. Among four flasks, the air from two flasks, after analysis for greenhouse gas mole fractions, was combined and analyzed for $\Delta(^{14}CO_2)$

Therefore we added the relevant sentence on Line131

Line144: Among four flasks, the air from two flasks, after analysis for greenhouse gas mole fractions, was combined and analyzed for $\Delta(^{14}CO_2)$

4. line 129 "suggested" ??

Line 144 we revised it to" tabulated"

5. line 245, 246 This largest positive C_{bio} is actually a single point. Which trajectory belongs to this point? The tic marks in the histogram of fig 2C do not correspond to those in the left part of fig 2C.

The largest positive value of C_{bio} is observed on August 2 2016 as shown in Figure 2. And it was observed in PL sector. To make it clear we revised the sentence line 279.

Line 287: The highest C_{bio} value was also observed in summer, PL sector.

Thank you for pointing out the y-axis ticks. We revised Figure 2 as below.

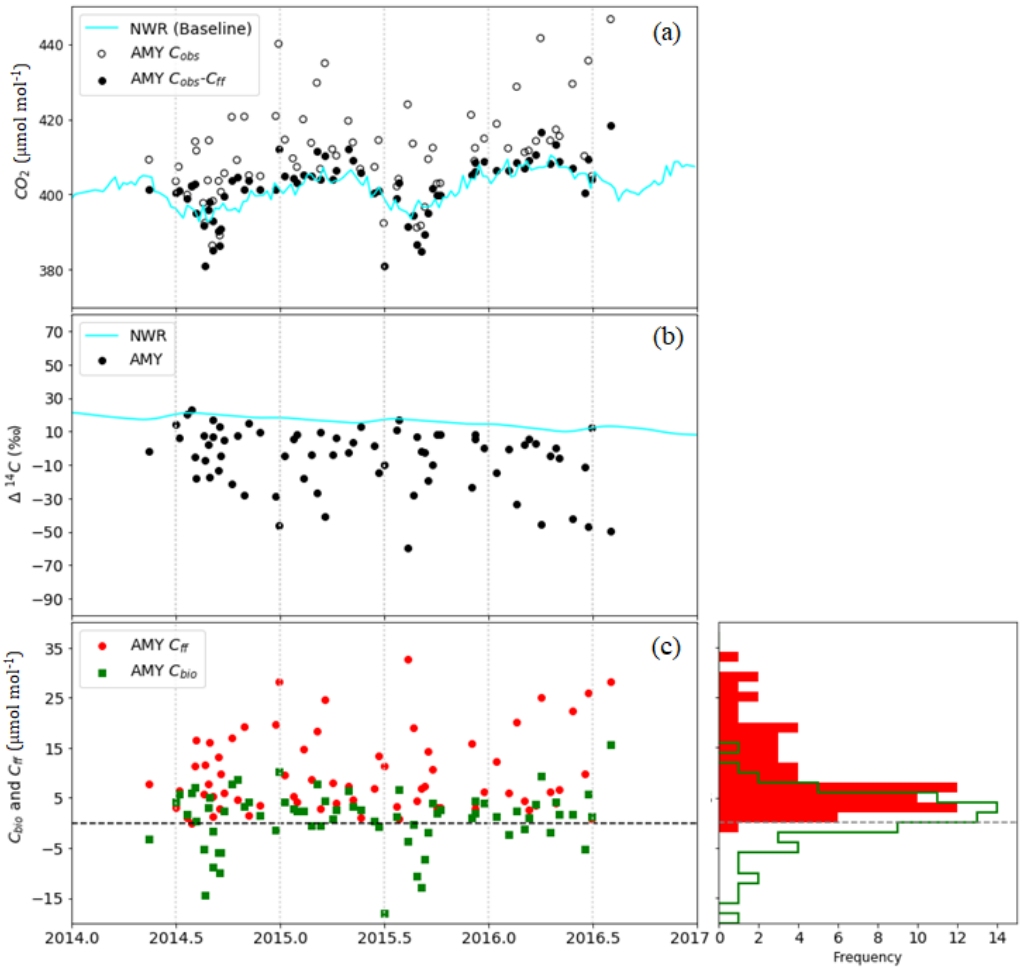

Figure 2. Time series of (a) observed CO2 dry air mole fraction (open circles) and observed CO_2 (C_{obs}) minus C_{ff} calculated from $\Delta(^{14}CO_2)$ (closed circles). (b) $\Delta(^{14}CO_2)$ at AMY (black circles) and at NWR (Niwot Ridge, line), baseline data. (c) Time series of C_{ff} and C_{bio} calculated from $\Delta(^{14}CO_2)$ (left) and the frequency distribution at AMY (right).

6. Lines 252-254. I doubt the explanation offered. Even though your sampling time is early afternoon, I expect still the average mixing height to be the most important player in the mixing ratios of Cff and Cbio, as it influences the flux-to-mixing ratios relation. There must be a seasonal effect in the mixing height no doubt. This must be taken into account in this discussion. The fact that Cbio covaries with Cff also points to the importance of the (average) mixing height.

We agreed with reviewer's comments. When we analyzed the PBL height for each sample by meteorological model, winter was highest (with a range from 150 m to 1100 m) and summer/spring are lower than other seasons (with a range from 100 m to 800 m). This result is consistent with the explanation of wind speed in the manuscript. Therefore we added this sentence.

Line 282: When we analyzed seasonal boundary layer height for each sample by UM-GDAPS, it also showed similar result that it was highest in winter (with a range from 150 m to 1100 m) and lowest in summer (with a range from 100 m to 500 m). This suggests that these high summer Cff values may reflect emission from local activities, which were described in section 2.1, more than in other seasons.

7. Line 270-271 I would say in general nobody expects the CO_2 enhancements above background to be entirely due to C_{ff}.

Thank you for the comment. The point is that regardless of the source and seasons, we find that C_{bio} contributes to atmospheric CO_2 enhancements at AMY. And if we just use CO_2 enhancements to demonstrate the bottom-up inventory, it can be biased. Therefore we emphasized this in section 4. "Finally, we stress that because C_{bio} contributes substantially to $\Delta x(CO_2)$, even in winter, 14C-based C_{ff} (and not $\Delta x(CO_2)$) is required for accurate calculation of both R_{CO} and R_{SF6}."

To avoid the confusion, we revised the explanation in the section directly on line 302

Line 302: … so when only CO_2 enhancements above background are compared to bottom-up inventories, it can make a bias due to C_{bio} contributions.

8. Line 280-282 "During the experimental period, the averages from Asian continent (sectors CE and CN) were higher than KL without the baseline level." Without the baseline level?? What do you mean?

“Without baseline” means that Continental Baseline (CB) and Ocean Baseline (OB) sector.

We added the explanation on line 316,

Line 316: …without baseline sector (CB and OB)

9. Does OB fit in this set? You call it ocean background, but at the same time you mention it crosses over Shanghai (213-214). In line 234-235 you indicate it again as being background, and then here (282) you take it along with the "real" continental trajectories. This needs to be clarified.

We agree. The cluster definitely belongs to Ocean Background, but some of the air masses (4 of 10) have paths over southern China, such as Shanghai. We confirmed that the altitudes that air masses came through were high enough when over southern China that they were unlikely affected by surface sources and sinks. But we cannot ignore the possibility that these trajectories could have affected our results. We revised the sentence below:

Line 239: .Among them, a few of the trajectories passed over the eastern part of China (e.g., over Shanghai) with high altitude (~1000 m).

10. 298-299 "we also see CN originated from northeast China and it was around (10.6±6.9) μmol mol–1." I don’t get the meaning or consequence of this sentence part.

A previous study mentioned that northeast China affected SDZ-NE sector samples and they showed (3±7) μmol mol^{-1} in 2009 to 2010 and increased to (7.6±6.8) μmol mol–1 in 2015.

In this study, samples originated from northeast China (NE) showed increased C_{ff} levels around (10.6±6.9) μmol compared to 2009 to 2010.

We revised the sentence clear

Line 333: we also see CN that originated from northeast china (NE) and its mean value of C_{ff} had increased around (10.6±6.9) μmol mol^{-1} compared to those values in 2009 to 2010.

11. Lines 300-302 Once more, I think average mixing height is the key player here. Are the weather patterns, and thus mixing heights different in the years 2009-2010 from 2014-2106? Did Turnbull et all also sample between 14 and 16 hours?

Turnbull et al. (2011a) sampled flasks during mid-afternoon, so that we assume the sampling time might be similar. One difference is they collected air when there was an onshore wind while in this study, samples were collected regardless of wind direction and speed.

We could not research the mixing height in 2009 to 2010, but it is very clear that C_{ff} from CB sector in this study increased compared to TAP far-field samples from 2004/2010. CB sector is the cleanest sector in this study with high wind speed (median value is 5 m/s, with a maximum of 10.2 m/s) and high PBL (median value is 600 m, with a maximum of 1700 m). There is no possibility for local-scale pollution to affect the samples due to the synoptic conditions.

Other sectors that originated from China, not only in this study, but also other studies, showed increased values compared to Turnbull et al. (2011a).

On the other hand, we also considered the possibility of unexpected conditions that could affect this analysis**. *Line 345: It is also likely that the proximity of local emission sources to AMY is causing higher observed C_{ff} under some synoptic conditions.***

As the reviewer mentioned, it would be great to mention those factors in more detail and the importance for further study, so we removed the sentence above and added:

Line 347: On the other hand, those values from this study showed large variability with small sample numbers due to different sampling strategy, environment, and synoptic

conditions such as boundary layer height at the sampling time from reference studies. Further study will be necessary to understand those increased values.

12. 304 increase of 16.7% line306 "broadly consistent" I disagree for the China case, as you find way larger increases between 2010 and 2016. So you might hypothesize that your measurements indicate much higher increases in fossil fuel use? Of course what you state in 311-312 is very true...

    We also pointed out the Asia main land differently from Korea Local air.

    Line 341: This is broadly consistent with the flat trend in observed C_{ff} in KL air masses, and in the upward trend in C_{ff} observed in air-masses flowing out from Asia.

    We also explained the possibilities to affect this result on line 347. Therefore we did not revise the sentence.

13. Line 321 are these differences significant? I would say (KL,PL) > (CN,CE) > (CB,OB)

    Line358: Corrected

14. Line 331 To my opinion SF_6 is not a good tracer/surrogate for fossil fuel CO2, as it is not produced in the same process. So SF_6 actually traces specific industrial activities, and electricity use. Both are coupled to fossil fuel CO_2, but not in a 1:1 (spatial, temporal) relation. CO, on the other hand is really co-produced with fossil fuel CO2 (and with biofuel CO2), albeit at a varying rate.

    Though the reviewer's comment is true, it is also true that C_{ff} and ΔSF_6 correlate quite strongly in CE, OB and KL sectors. This means even if they don't have a same source, they are emitted from similar regions. Therefore, we suggest that ΔSF_6 can be a proxy of fossil fuel CO_2 in those regions. But as you mentioned ΔSF_6 cannot always a good tracer of fossil fuel CO_2, so we revised the sentence as below.

We revised line 362:

Line 372: Thus SF_6 can be a good tracer of fossil fuel CO_2 **for those regions**.

15. Line 332 "Even though" I don't see the contrast between the strong correlation and the differences

The correlation was strong in both South Korea and for the Asian continent, but the R value was totally different for both regions.

We revised line 373:

Line 373: The correlation between $\Delta x(SF_6)$ and Cff was strong in CE, OB and KL, however, R_{SF6} is different between South Korea and outflow from the Asian continent (Figure S2).

16. Caption figure S3: "From 2005,.. " -> "From 2005 onwards, .."

Corrected.

17. lines 347-349 Still, in spite of the still large uncertainty, I invite you to make a stronger statement here, namely that the SF_6 inventory in EDGAR and in KNIR are too low given your measurements.

We revised the sentence here

Line 388: Even though KL R_{SF6} showed greater uncertainty than CE and OB, it is still greater than bottom-up inventories, such as KNIR and EDGAR. Therefore it would be useful to get more data to try and derive a more robust estimate to evaluate SF_6 emission inventories for Korea.

18. Line 351 also here, watch the significance. I would conclude from table 1 that CB=KL. And indeed

(see my point higher up), OB is mostly regional background air.

Corrected (Line 395).

For OB sector, we revised the explanation according to reviewer's comment No.9 above.

19. 354 "CO...it is more closely related to fossil fuel CO2 emissions" yes, but also to biomaterial combustion (compare the Cff to the CO excess)

We agree. The revised sentence below

Line 397: Since CO is produced during incomplete combustion of fossil fuels and biomass, it is more closely related to fossil fuel CO_2 emissions than the other trace gases.

20. 357-358 I think you can safely erase the word "likely" here. 358 add "and the use of catalysers" ?

Corrected

Line 402 … due to differences in combustion efficiencies and the use of catalytic converters

21. 358-360 Indeed, biomaterial combustion must play a role, regarding the low C_{ff} especially for CB.

We agree and added the following sentence.

Line 405: For example, for CB the CO level is similar to KL while R_{CO} is higher than KL with low Cff.

22. 366 Figure S2 -> Figure S1

Corrected (**Line 413**).

23. 369 Paragraph 3.4 I suppose you did a similar thing for the SF6 inventories. That means the either the start of this paragraph should be moved up into 3.3, or the SF6 inventory discussions should be taken form 3.3 and moved to this paragraph.

Corrected. We separated two sections according to the species.

3.3 Correlation of Cff with SF6 and its emission ratios

3.4 Correlation of Cff with CO and its emission ratios

24. 377 "The uncertainty of EDGAR4.3.2 emissions" -> "The uncertainty of EDGAR4.3.2 fossil fuel CO2 emissions"

Corrected (**Line 422**)

25. 397-399 if a difference is not significant, it is doubtful to discuss its possible causes.

We removed the sentence from Line 443 to Line 444.

26. 403 "KNIR seems to have uncounted CO emissions," -> "KNIR suffers from a high number of missing CO emission sources," in other words: make this statement stronger, as the difference is huge: _2500 vs _700 Gg in 2012 ! And your data corroborate the Edgar emission ratios...

Corrected

Line 448: KNIR suffers from a large number of missing CO emission sources compared to the EDGAR, as indicated by their reported emissions, 638.3 and 2580.8 Gg in 2012, respectively

27. 433-439 S. Korea: your RCO results are 1.2 times the Edgar results. That is hard to see in figure 4.

Your value (from table 1) is 8±2, so a ±25% uncertainty, which makes this factor 1.2 not significant. The Chinese inventories, on the other hand, ARE significantly too low, even though the declining trend has been confirmed by atmospheric measurements. My guess would be that the lack of biofuels/biomaterial burning which is not present in the EDGAR CO inventory, explains the large difference in China, and is not so important in S. Korea.

We agree reviewer's comment for two reasons. Frist, EDGAR does not reflect secondary CO production and, second, CO derived from biomass burning and biofuels was not included by EDGAR. Since we described it already, we did not revise the sentence. This is described in the manuscript as:

Line 485: Also, CO derived from biomass burning and biofuels was not included in this inventory. Therefore, this indicates that top-down observations are necessary to evaluate and improve bottom-up emission products.

28. 441 (and also earlier and further) you express mean values ± standard deviations, whereas the way you write it suggests that this is the error in the mean value, which is in fact sqrt(#mm) lower. So in fact the mean value here is (-6.2±2.2) ‰ (I took N=70), with a spread of 19‰. In your case, most of the time the spread= the standard deviation is the important feature, but if you compare in lines 446-447 to previous measurements at TAP it is important to know how many measurements those were, and thus what the mean and error in the mean are. Your statement: the average is twice as high strongly suggests that this difference is significant, but the reader can only judge that if you present the error in the mean in both cases. I advise to make this difference between standard deviation and error in the mean clear at the various points where it matters in the paper.

Thank you for your comment. It is very true that standard deviation (SD) and standard error of the mean (SEM) are totally different. The reference values from Turnbull et al. (2011a) and Niu et al. (2016) were also SD rather than SEM, which we verified in their publications.

As the reviewer already commented, SD is the dispersion of data in normal distributions. In other words, SD indicates how accurately the mean represents sample data. For SEM, it is the SD of the

theoretical distribution of the sample mean (the sampling distribution).

In this regards, it would be good to use SEM though, however, since the data set we used here is not continuous, only weekly resolution, the number of data is quite small for SEM. When we use SEM, the data characteristics can be underestimated because the error can be decreased, certainly. The number of data is very limited that dispersion of data can be important information for reader.

Therefore we just added the number of the data for the calculation (since whether we use SD or SEM, this is very important information) and retain SD.

Please see the revised manuscript. When previous studies did not include the number of data, we could not include it.

29. Lines 449-452 Yes, Cff really increased for the air masses from the Asian mainland. Do you conclude that this indicates stronger growth of fossil fuel use than the statistics say? If you think your data clearly point at that, mention that here.

The atmospheric Cff was increased compared to the previous studies but we cannot explain that our results are much greater than the reported inventory values. When we analyze inverse modeling, we can point it clearly. Therefore we revised sentence in section 4, 2).

Line 502: After separately identifying samples originating from the Asian continent and the Korean peninsula, we determined that the mean C_{ff} increased relative to the earlier observations due to increased fossil fuel emissions from the Asian continent as showing the consistent growth with reported emission increased 16.7% in China while 1.8% in South Korea from 2010 to 2016.

30. lines 453-463 Based on your data I would (also) conclude the following: (1) 14C analysis is a reliable way of determining Cff in the mixing ratio of air masses (2) Then, the ratio of the emission of rare trace gases and Cff can be determined as well (3) As the inventories for various other trace gases/greenhouse gases are generally much less reliable than that of Cff, these inventories can

be validated/verified using atmospheric measurements like ours. (4) I our case we conclude that the inventories for SF6 ... and for CO ... In this way your results will probably be more valuable to policy makers. would also formulate (part of) this reasoning in the abstract.

We re-write the summary and conclusion according to reviewer's comment.

Please see the revised the version of section 4 and abstract as well.

31. Two more references suggested: Page 2 I would suggest in addition the reference : van der Laan, S. et al. Observation based estimates of fossil fuel-derived CO2 emissions in the Netherlands using Delta 14C, CO and 222Radon, Tellus B, 62(5, SI), 389–402, doi:10.1111/j.1600-0889.2010.00493.x, 2010.

We added the reference.

32. page 3 line 64 "..correlate well..." I think the earliest 14C-based reference to this is Zondervan, A. and Meijer, H. A. J.: Isotopic characterisation of CO2 sources during regional pollution events using isotopic and radiocarbon analysis, TELLUS SERIES BCHEMICAL AND PHYSICAL METEOROLOGY, 48(4), 601–612, doi:10.1034/j.1600-0889.1996.00013.x, 1996.

We added the reference.