

Answer to Referee 1

General comments:

The paper is not very well organized and seems to have been put together from disparate pieces. Some parts that should be in the Results section are instead in the Methods and vice-versa (see specific comments).

The comparison with the WMO recommendations: in the abstract is stated: “the GC system meets the WMO recommendations” (and this is said several times through the paper). Comparing the reported numbers, e.g. Table 6, with the WMO recommendations (page 3140, lines 23-24), it looks like this statement is not true. Please clarify what do the WMO recommendations refer to, what should be compared to what, and please check again the comparison and change the text as need. See also the specific comments.

R: We modified the specified parts in order to be more precise regarding the WMO recommendations.

Puy de Dome is a mountain station. Is there no special atmospheric circulation or behavior due to the fact that this is a mountain station? (see e.g. Thompson et al., 2009) This does not seem to be taken into account, for example when comparing to the PBL height.

R: There are definitively specific patterns in the atmospheric transport processes at a mountain site, like Puy de Dôme. This is the reason why we select only night time mole fractions in order to represent large scale processes (see e.g. the discussion concerning the marine boundary layer reference), as opposed to stations located in flat terrain where the background data selection generally target only mid-afternoon measurements. The discussion of our results refers to the typical variabilities observed at similar mountain sites like Schauinsland, and Kasprowy.

Specific comments:

1. page 3123, lines 9 – 11: the phrase “ Comparisons ... the GC system meets the WMO recommendations” – this doesn’t seem to be true for all species at all times, please clarify.

R: we removed this sentence from the abstract and gave more explanations on the comparison later in the manuscript

2. page 3123, lines 20 – 22: I suggest to give the errors of the fluxes

R: the respective errors have been added

3. page 3125, line 2: I think “These approaches” should be “Some of these approaches”

R: corrected

4. page 3125, line 6: Lopez et al., 2013 did not use radiocarbon in CO₂ – please reformulate

R: we did use radiocarbon in CO₂ in the following paper:

Lopez, M., Schmidt, M., Delmotte, M., Colomb, A., Gros, V., Janssen, C., & Bousquet, P. (2013). CO, NO_x and ¹³CO₂ as tracers for fossil fuel CO₂: results from a pilot study in Paris during winter 2010. *Atmospheric Chemistry and Physics*, 13(15), 7343-7358.

5. page 3126, line 5: remove “can be cited”

R: done

6. page 3126, line 7: “achieve equivalent or superior repeatability...” – please give references

R: we added the following references

- Hammer et al. 2013

- Yver Kwok et al., 2015.

7. page 3126, line 16: “anthropic” should be “anthropogenic”. Please check the whole document, this appears several times.

R: this has been corrected for the whole manuscript

8. page 3126, line 16: “SF6 ...is the forth ...GHG in terms of radiative forcing” – please give reference

R: we re-arranged the sentence:

“which is an extremely stable GHG having a global warming potential of 23,900 (Forster et al., 2007)”

9. page 3127, lines 5 – 6: “which demonstrated that our measurement system reaches the WMO recommendations” – please check how much of this is true, and adapt text

R: we adapted the text by changing the sentence.

“A paragraph is dedicated to data quality control and atmospheric measurement comparisons showing in which conditions our measurement system reaches the WMO recommendations”

10. page 3127, lines 18 – 19: “is surrounded by meadows (36.4 %), forests (33.4 %) and arable land (17.6 %)” - in which area or radius around the station?

R: these values are given for the Auvergne region, Puy de Dôme station being in the center of it. The region having a square shape, we then corrected the sentence by specifying that the given values are for the region. The following sentence has been added:

“the ground cover of Auvergne region (26 013 km²) consists mainly of meadows (36.4%), forests (33.4%) and arable land (17.6%), Puy de Dôme station being located in the center of this region”

11. page 3128, line 25: give the coordinates in the same units as for the station (see page 3127 line 13)

R: corrected

12. page 3128, line 25: why were the ECMWF data extracted at 575 masl and not at the station altitude?

R: The main purpose of this extraction was to retrieve the PBL height. The volcano on which the station is located can be considered as a high singularity in a plateau of approximately 575m asl.

13. page 3129, line 15: what was released, air masses or particles?

R: We adapted the sentence:

“Eight particles were released every 15 min (96 particles every 3 h) in a 3-D box”

14. page 3129, line 19: “particles arriving at the station” should be “air masses arriving ...”?

R: particles have been released

15. page 3133, line 4: not sure whether “CO2 catalysis” is the right term – should it be “CO2 reduction over the Ni catalyst”?

R: corrected

16. page 3133, line 9: “valve shunting” – should it be “valve switching”?

R: corrected

17. page 3133, line 15: “ECPs” should be “EPCs”

R: corrected

18. page 3134, lines 1- 3: for calibration, the whole air samples are compared to synthetic air-based calibration gas, which is missing many components normally present in atmospheric air (e.g. He, H₂, CO, short lived gases). Is it known that this has no influence on the calibration? I.e. has it been proven that none of the gases present in atmospheric air has any influence on the measurement? In the same idea, please specify whether the calibration cylinders for the CRDS are also based on synthetic air. If not, could this be (part of) the cause of the difference between the GC and CRDs?

R: We are aware, that it would be better to use real air calibration gases, but the spiking for 4 components is quite difficult. We have compared synthetic and air based tanks and have not seen differences or problems.

The calibration cylinders for the CRDS are also based on synthetic air. They have been calibrated by a GC (which see ¹²CO₂ and ¹³CO₂) which indeed induce a maximum bias of 0.07 μmol mol⁻¹ in the CRDS measurement. The ambient measurement from the CRDS will be corrected for this bias but it does not explain the 0.25 μmol mol⁻¹ offset.

19. page 3135, line 1: “a linear fit is sufficient to account for the non-linearity of our uECD”; also, line 4: “a two-point calibration strategy is well adapted to correct for the uECD non-linearity”. The linear fit by definition does not account for any non-linearity. Also, a two-point calibration cannot correct for any non-linearity. I think what the authors wanted to say is that the non-linearity is not corrected, and the error due to not accounting for non-linearity over this range is small / acceptable within errors. Please reformulate.

R: We changed the paragraph, which is now as follow:

“The response function of μECD for N₂O analysis is non-linear, especially in the range below and above the tropospheric values (see Schmidt et al., 2001; van der Laan et al., 2009b; Lopez et al., 2012). The non-linearity of our μECD was tested by analyzing five cylinders calibrated by the CCL on the NOAA-2006A scale and with N₂O mole fractions between 302.00 and 338.04 nmol mol⁻¹. In this small mole fraction range, which is important for our measurements, a two point calibration describes sufficient the response function of our μECD.”

20. page 3135, line 11: what is the starting gas pressure in working standards?

R: we added in the manuscript that the starting gas pressure is 200bar:

“These gas mixtures are used to fill 40L aluminum cylinders (Luxfer) to 200bar”

21. page 3135, line 15: the initial and final analysis of these cylinders at LSCE showed significant differences, at least for CO₂ and N₂O. Please specify: are these differences statistically significant? (give errors?); which of the values were used for calibrating the Puy de Dome data? what is the known / suspected cause of this difference, drift in the cylinders (did both standards show similar differences?), LSCE lab calibration or measurement stability, other? If the cylinders are suspected to have drifted, how is this taken into account for the calibration at Puy de Dome? If the LSCE lab is the problem, which data are considered “true”?

R: For N₂O, the differences are not statistically significant as the calibration precision is in the order of 0.15 nmol mol⁻¹. For the CO₂, we indeed observed a small drift coming from the calibration cylinders (-0.08 and -0.15 μmol mol⁻¹), this drift being also observed in the TGT (fig.3). As the drift is about

0.03 $\mu\text{mol mol}^{-1} \text{yr}^{-1}$, we did not have corrected this drift. We added the following sentence:
“The observed differences are not statistically significant but for CO₂. The calibration cylinders have respectively drifted by 0.08 and 0.15 $\mu\text{mol mol}^{-1}$ over their lifetime for CO₂. The CO₂ data presented in this paper are not corrected for the observed drift on the order of 0.03 $\mu\text{mol mol}^{-1} \text{yr}^{-1}$.”

22. page 3135, line 26: the tables are not numbered in order. The last was Table 3, so this one should be Table 4.

- the same at page 3136, line 11

- change also the name and the order of the tables at the end of the document

R: we changed the table numbering.

23. page 3135, from line 27: the text from here till the end of this subsection belongs to the results and discussion

R: We made a distinction between the signals observed in ambient air which are discussed in the results section, and the evaluation of the instrument uncertainties based on target gas measurements which is presented in this section dedicated to the description of the GC analyzer.

24. page 3136, line 1: “data holes” – better “data gaps”?

R: corrected

25. page 3136, lines 5, 6: “SD at 1-sigma” – SD (standard deviation) is sigma. Remove “at 1-sigma”

R: corrected

26. page 3136, line 7: “Table 4” should be “Table 5”

R: corrected

27. page 3136, lines 14-17: I cannot understand the phrase starting with “The same ...” – please consider reformulating.

R: We changed the sentence to: “The respective differences between the CO₂ assigned and measured values of the TGT gas in period A and period B were 0.15 and 0.19 $\mu\text{mol mol}^{-1}$ confirming the consistency between the two scales used in periods A and B. This agreement when using two different calibration scales shows that the problem is probably due to an error in the value attributed to this target gas.

28. page 3136, lines 22-24: Were the data corrected, or the calibration was just re-done with one standard instead of two? More explanations are needed on how the correction was done. If the data including Target were corrected or recalibrated, why is the 10 ppb difference still there in Fig. 3?

R: the data have not been corrected yet. We are still developing for new database tools in order to take into account a one-point calibration strategy in case of what we observed. We corrected the sentence: “Target data as well as ambient air measurements for this period should be re-calibrated by applying a one point calibration to the four trace gases”

29. page 3136, lines 25-27: The phrase starting with “The analysis of TG2 ...” – these data after Sep 2013 are not shown in the paper; also the WMO recommendations are not specified. I suggest to remove this phrase, or support it with data.

R: we removed this sentence as the paper focus only on the July 2010 to July 2013 period.

30. page 3137, line 7: “as the GC” should be “as the GC trap”?

R: corrected

31. page 3137, lines 25-27: what are the similarities between the two stations that are relevant here? (technical setup? atmospheric circulation? other?)

R: we specified that the two stations are two medium-elevation mountain sites and have the same geographical environment . They are frequently above the continental boundary layer, especially in winter.

32. page 3138, lines 17-18: here it is said that the reference cylinder is analyzed for 10 min every hour; but the CO₂ analyzer is differential and has a reference cell that should be flushed continuously. Or is the reference cylinder not the gas that goes through the reference cell? Please clarify.

R: The reference cell of the instrument is continuously flushed with a reference gas. This gas is passed through both cells for 10 min every hour. This has been clarified in the manuscript in the following sentences:

“The measurement frequency is 1 Hz , and the cell flow is typically 20 mL min⁻¹ for the sampling cell, and 15 mL min⁻¹ for the reference cell which is continuously flushed with a reference gas.”

“Ambient air is analyzed for 50 min following the analysis of the reference cylinder for 10 min , which is passed through both cells at the same time.”

33. pages 3139 – 3141, Sect. 3.4: Most of this section, except for the first paragraph, should be in the Results. Some additional information on the comparison method should be given here in the method, e.g.: how were instruments compared, since they have very different data frequencies? The CRDS works at 1 Hz, the GC gives 1 data point every 5 min, and flask sampling is weekly – they cannot be compared without some data processing.

R: We gather 3.2 and 3.4 paragraphs now called “Quality control of the GC system and comparisons with different analyzers”. As explained in comment 23, we prefer keeping this part in the section dedicated to the description of our measurement system.

For the continuous measurement comparisons, we added the following sentence in the manuscript: “These differences are calculated from the hourly mean measurements”

For the continuous vs flask comparison, we added the following sentence in the manuscript: “In order to match the flask sampling time, the differences between GC measurements and flask measurements have been calculated using a linear interpolation on the mole fractions corresponding the two GC measurements bracketing the sampling time of each sampled flask.”

34. page 3140, lines 1 – 20: were any of the station gases (reference or target) measured by both GC and CRDS? If yes, was the difference similar with the ambient air difference?

R: We indeed measured the second set of the GC calibration cylinders on the CRDS. The difference between the attributed values (at LSCE) and the measured mole fraction by the CRDS (at Puy de Dôme) were 0.03 (for the WH) and 0.34 μmol mol⁻¹ for the WH and WL, respectively.

It could explain the differences observed in the ambient measurements, but the ambient air differences (CRDS and GC) plotted against the GC mole fraction is constant. To clarify, we added the following sentence in the manuscript:

“This observed difference is stable over time and is not concentration dependent. A second experiment was to analyze the second set of working standard on the CRDS. The results showed a difference between the assigned value (at LSCE) and the CRDS of 0.03 μmol mol⁻¹ on the WL and of 0.34 μmol mol⁻¹ on the WH.”

35. page 3140, lines 1 - 20: The CO₂ differences between instruments listed here (and in Table 6) are obviously not within the WMO targets. Please check the affirmation made through the paper on this

subject.

- Similar for lines 23-24: “the comparisons between GC in situ measurements and flask analyses reached the desirable comparison levels (see Table 6 for more details)” – in Table 6 it is shown that CO₂ and N₂O (in-situ – flasks) are not within the WMO targets. Please adjust text or, if “desirable” in this case is not WMO, please specify.

R: We added the following sentences in order to clarify:

“The different comparison methods presented in this section showed that the GC system installed at Puy de Dôme station matches the WMO-GAW recommendations for CH₄ and SF₆ measurements. The CO₂ comparison shows different results depending on the method used. The recommendations are reached by considering the comparison with cylinders or flasks whereas it is not if we consider only the comparison between in-situ instruments. The WMO-GAW recommendations concerning the N₂O measurements are ambitious considering the repeatability obtained with our GC. The N₂O measurements do not reach the WMO-GAW recommendation but we can notice that the different comparison methods used show differences lower than our instrumental repeatability.”

36. page 3141, line 20: 1013 should be 2013

R: corrected

37. page 3142, line 17: remove “than in winter” or move to the end

R: “than in winter” has been removed

38. page 3143, lines 14-15: how was the time interval for the “background” chosen? Looking at Fig. 5, there are quite some differences between the start and the end of this interval, probably most evident in Rn and CH₄. The continental offset and thus the computed fluxes depend on this choice

R: To avoid the influence from the residual layer and from the closed surface emission, we chose to use measurement performed in the free troposphere. This choice has been made relative to the diurnal cycle of the PBL height presented in fig5. This choice is also discussed in the answer 44 (please see below).

39. page 3144, Sect 4.3.1: I think most of this should sit in the Methods.

R: there is no Methods part. For the understanding and the flow of the paper, we would prefer not moving this part.

40. page 3145, lines 1-9: Schmidt et al., 2003 calculated the decay correction based on the 3 days continental residence time for the air masses at Schauinsland. I do not understand how the fact that “most of the air masses arriving at the Puy de Dôme station and having a lifetime of three days are also from the continent” proves that the air arriving at Puy de Dome has a continental age of 3 days. In short: if all the 3-days air is from the continent, it does not mean that all the air is 3-days old. Please clarify.

R: We removed the concerned sentence and we clarified by adding the following sentences:

“The term in brackets in eq. 1 corresponds to the radioactive decay correction factor. For a continental mountain site like Schauinsland, Schmidt et al. (2003) determined that the mean residence time for air masses over the European continent before reaching the station is between 2 and 4 days, leading to a net effect of radioactive decay varying from 16% to 29%, respectively.”

41. page 3145, lines 6-7: Please give the Schauinsland coordinates. Is there any study that compared, and showed that the two stations are similar? (give reference)

R: Schauinsland coordinates are given previously in part 3.3.2.

Study of Lukacs et al., 2007: show the similarities of ‘brown carbon’ (defined as light-absorbing organic matter) and brown carbon/²¹⁰Pb (²¹⁰Pb being the daughter element of ²²²Rn) measurements between the two sites in term of seasonal cycle, trend and transport.

Study of Hammer et al., 2007: showed based on ²¹⁰Pb measurements at PUY and SIL that the regional-scale circulation patterns are frequently common to both sites.

We added these two references in the manuscript.

42. page 3145, line16: why was the ²²²Rn not extracted for the station footprint? How was this 300 x 300 Km area decided?

R: the Rn emission map is published now (Karstens et al., 2015), we have updated the manuscript in consequence. We chose to focus on the 9 pixels contributing the most to the signal measured at the station during nighttime based on the footprint presented in fig1b. Also, the selected region is the highest potential sources of radon in France due to its geological formation:

[http://www.irsn.fr/FR/connaissances/Mediatheque/pages/Detail_Mediatheque.aspx?](http://www.irsn.fr/FR/connaissances/Mediatheque/pages/Detail_Mediatheque.aspx?GuidWeb=2c2a9274-9106-41cf-b110-445981d4784e&GuidList=46306625-d53e-464d-9750-4113c34c8065&GuidItem=18&imgId=467a23db-0d7e-484f-8890-da26f97c175e|43&Cible=1)

[GuidWeb=2c2a9274-9106-41cf-b110-445981d4784e&GuidList=46306625-d53e-464d-9750-4113c34c8065&GuidItem=18&imgId=467a23db-0d7e-484f-8890-da26f97c175e|43&Cible=1.](http://www.irsn.fr/FR/connaissances/Mediatheque/pages/Detail_Mediatheque.aspx?GuidWeb=2c2a9274-9106-41cf-b110-445981d4784e&GuidList=46306625-d53e-464d-9750-4113c34c8065&GuidItem=18&imgId=467a23db-0d7e-484f-8890-da26f97c175e|43&Cible=1)

The signal coming from the rest of France is low and contributes weakly compared to the selected area.

43. page 3145 line 28: typo “therefor”

R: corrected

44. page 3146, lines 6-8: Only random errors were considered here (GC repeatabilities). What about possible systematic errors, e.g. due to imperfect choice of background time intervals?

R: We did a study case consisting of calculating the background from 22h00 to 06h00 (based on the results from PBL height), 00h00 to 03h00 and finally from 03h00 to 06h00. Results are given in the following table. The average fluxes are consistent for the CH₄ and the N₂O, given the uncertainties in order of 60%. They are not consistent for the CO₂, pointing out that the background is indeed a key factor for the flux calculation. We would prefer to use an average over the whole night, when the station is in the free troposphere as there is no strong argument in favor of choosing a particular time slot overnight.

fluxes	22h00 – 06h00 (actual values)	00h00 - 03h00	03h00 - 06h00
CO ₂ (t km ⁻² yr ⁻¹)	1150	1600	-225
CH ₄ (t km ⁻² yr ⁻¹)	5.6	5.1	6.0
N ₂ O (kg km ⁻² yr ⁻¹)	1546	1530	1493

45. page 3148, lines 4-5: Is there an explanation for the negative CH₄ fluxes? From Fig. 4, it seems that for the negative fluxes – months there were fewer measurements. Can this be the cause, that the baseline is not well defined during these months?

R: Lower background values (compared to the MBLR) occur during periods with many data gaps which can bias the observed background. Air masses arriving from higher altitudes also induce lower background. We added the following sentence in the manuscript to clarify:

“ Negative values occurred in April, September and November 2011 due to biases in the calculated background induced by the many data gaps in the considered months. Therefore, these negative fluxes are not taken into account in the average flux calculation.”

46. page 3148, line 29: I do not get the meaning of “ with lower fluxes as expected” – please check.

R: we removed this part.

47. page 3149, lines 4 – 7: Please check the comparison with the WMO targets. From the flask comparison, it seems that the CO₂ and N₂O are not within these targets; for the cylinder comparison, N₂O does not match.

R: we added to following sentences in the conclusion part:

“We described in details three methods which have been used for comparison purposes. They are based on direct comparison between two in-situ analyzers, flask measurements and cylinder measurements. For CH₄ and SF₆, all comparisons show that GC measurements at Puy de Dôme are in agreement with the WMO-GAW compatibility goals. For N₂O, our measurements do not match the ambitious WMO-GAW compatibility goal. For CO₂, the comparison based on ambient air flasks and reference cylinders analysis between the GCs operated at Puy de Dôme and at LSCE reaches the desirable comparison level, showing there is no bias in the scale transfer between the two sites. Nevertheless, it does not for the in-situ comparison with other analyzers (NDIR and CRDS). The comparisons between the GC and the CRDS in-situ measurements indicated a constant offset of 0.21 μmol mol⁻¹ CO₂ over 20 months of overlapping measurements ”

48. page 3149, lines 12-14: phrase starting with “ At stations...” : the discussion till here was about the difference between the two analyzers. At stations with only one analyzer, there can be no bias between analyzers.

R: we added a line break and added the following sentence:

“For consistency reasons, we thus recommend to use different methods based on flask or cylinder comparisons but also based on in-situ comparisons to check if the considered measurements match with the WMO-GAW compatibility goals.”