Response to Jocelyn Turnbull's interactive comment on "Estimation of continuous anthropogenic CO_2 using CO_2, CO, $\delta^{13}C(CO_2)$ and $\Delta^{14}C(CO_2)$"

We want to thank Jocelyn Turnbull for her very helpful comments and suggestions. We have revised the manuscript and describe the changes in the following.

General comment:
This paper describes a modelling study that tests how well different tracers (CO2 mole fraction, CO mole fraction, δ13C of CO2 and Δ14C of CO2 and combinations thereof) can be used to determine recently added fuel CO2 mole fractions in the atmosphere.
The study is well designed and comprehensive, and the results are both topical and very useful. The results show that in Europe, CO may not be as useful a tracer for fuel CO2 as in other regions, owing to the low CO:CO2 emission ratio from European traffic – a result that is frustrating from a detection point of view, but satisfying from a clean air perspective. The study shows that δ13C of CO2 can be quite useful for quantifying fuel CO2, as long as the isotopic content of the various sources and the relative source mix is well known. An extremely helpful outcome (especially for ICOS) is the test of how best to utilise 14C measurements to "calibrate" the CO and δ13C methods, including both how many and what type of 14C samples are realistically needed to provide robust calibration of the other methods.
The paper is quite long, and this is understandable since there is a lot of information to be drawn from the model study. The decision to include all of it in one paper seems sensible since there's no obvious place to split it into a second paper. However, the discussion and conclusions sections are somewhat repetitive of one another and of the results section, so they could be condensed to reduce the overall length. Overall, this paper is well worthy for publication in ACP, with some minor revisions.

We agree that the discussion and conclusion is quite repetitive and have significantly shortened and combined the discussion and conclusion.

Specific comments:

The title could be improved to better reflect the subject. Something like - Continuous estimation of anthropogenic CO2: model-based evaluation of CO2, CO, δ13C (CO2) and Δ14C(CO2) methods.

We have changed the manuscript title according to the suggestion to: Continuous estimation of anthropogenic CO2: Model-based evaluation of CO2, CO, δ13C (CO2) and Δ14C(CO2) tracer methods

Pg 20183 line 17-19. Is there a particular basis for reducing fuel CO2 uncertainties by half (vs reducing uncertainties even further)? For example, what anthropogenic CO2 uncertainty is needed in global modelling studies so that this uncertainty becomes inconsequential relative to other uncertainties and biases in the models? What about for urban studies?

The required fuel CO2 uncertainties depend on the site investigated. As the reviewer implies, in urban studies the required absolute uncertainty of fuel CO_2 contribution is smaller than in rural sites since the signal (and therewith the signal to noise ratio) is higher. We therefore agree, that halving the fuel CO2 uncertainties should not be stated as universal goal. In this study, we investigate how well (in terms of bias and precision) we can estimate continuous fuel CO2 and leave it to the method-applicant to decide which accuracy or precision is sufficient for his/her purpose.
In the discussion of our revised manuscript we also add the effect of the model transport error on the emission estimates as retrieved from inverse models. This seems to be in the order of 40 % for afternoon values (e.g. Gerbig et al., 2008) and thus a major factor limiting the precision of top-down emission inventories next to the uncertainty of continuous fuel CO2 estimation.

Intro first paragraph. Here the concept of "fuel CO2" is introduced, and the majority of the paper talks about how to constrain the total fuel CO2 (fossil + biofuel). Only very late in the paper is it mentioned that it may also be useful to determine the fossil fuel CO2 component separately. If the objective of these measurements is to constrain fossil fuel emissions – which is critical to ensuring that emission regulations are working - then total fuel CO2 is not terribly useful. It would be helpful to bring this up early in the introduction, and explain why the focus here is on total fuel CO2 (presumably because in large scale models one objective is to solve for the biospheric flux, requiring the total fuel flux to be known).

In the revised manuscript, we discuss and elaborate the concept of fuel CO2 and fossil fuel CO2 already in the introduction and note that for some applications (e.g. determining the non-regenerative fuel CO2 contribution) it is worthwhile to estimate fossil fuel CO2 rather than fuel CO2.

Page 20184 line 7-8. Miller 2012 does not attempt to monitor fossil fuel CO2 emissions, rather they assume that these emissions are known and use them to examine emissions of other species. There are many more appropriate references that could be used here.

We have corrected this in the revised manuscript and used more appropriate references (Levin et al., 2003; Turnbull et al., 2006; Levin and Karstens, 2007; Turnbull et al., 2015).

Page 20184 line 10. Accelerator mass spectrometry!

We have corrected this.

Page 20184 line 9-12. The main issue with 14C measurement is the need to collect discrete samples, limiting the frequency. The 14C measurement cost is not prohibitive compared to the expensive instrumentation used for the other species.

In the revised manuscript we emphasize that the main limitation of 14C measurements is the measurement frequency. However, also the costs of 14C measurements should not be neglected if taking into account full costs, which are of order of 300 – 500 Euros per sample, depending on precision. We still emphasize that optical instruments are able to measure multiple species simultaneously with high temporal resolution, lower investment costs and much lower running costs.

Page 20184 lines 12-19. The GC-AMS system referred to here (McIntyre et al 2013) does not appear to suitable for continuous measurement of 14C in atmospheric CO2.
It is a lab-based method that is not field deployable, so there is no obvious way that it could be used for continuous atmospheric 14CO2 measurement, nor is it designed to separate CO2 from air. The 6‰ uncertainty on modern samples determined in that paper requires 0.67 mgC, somewhat larger amounts than are currently used for flask/graphite/AMS measurement of 14CO2 (cf Graven et al 2007; Turnbull et al 2007), and it requires multiple injections to achieve this uncertainty which likely takes as long or longer than measurement of a single graphite sample. It is likely that laser-based 14CO2 measurement systems will become possible in the next few years, and at some point these may be field deployable for continuous measurement. Even once the significant technical challenges of these methods are overcome, it appears that precision will initially be much poorer than AMS or gas counting, on the order of a few percent precision, even with time averaging. Thus for this modelling study, it would be more reasonable to consider the possibilities of (a) high resolution flask 14CO2 measurements at 2‰ precision, (b) hypothetical laser-based 14CO2 measurement with 3% precision, (c) use of flask or gas counting 14CO2 to "calibrate" the other methods, as is already discussed in the paper.

We acknowledge that the GC-AMS system described in McIntyre et al., 2013 cannot be used for continuous measurement of atmospheric 14C(CO2) as such. However, the reason why we have chosen a precision of 5 ‰ is that we want to investigate if future available continuous measurements of 14C(CO2) with such a precision may help us to determine continuous fuel CO2 in future. Since we do not know what precision of continuous 14C(CO2) measurements will be reached, we have chosen a number of 5 ‰, which we assume to be at the high end of what is possible. If 14C(CO2) measurements with this precision could help us to determine fuel CO2 effectively, it would highlight the benefit and usefulness of developing such continuous 14C(CO2) measurements. We therefore would like to keep the high estimate of 14C(CO2). However, in the revised manuscript we point out that this is a high estimate and we additionally note in the text what the precision would be when using 1% and 2%, as may be achieved in the near or middle-term future by laser-based 14CO2 measurements (Galli et al., 2013).

Pg 20185 lines 10-15. The authors may also want to refer to the following paper, which uses the combined 14C/CO method proposed by Levin and Karstens 2007. Turnbull, J.C., Karion, A., Fischer, M.L., Faloona, I., Guilderson, T., Lehman, S.J., Miller, B.R., Miller, J.B., Montzka, S., Sherwood, T., Saripalli, S., Sweeney, C., Tans, P.P., 2011. Assessment of fossil fuel carbon dioxide and other anthropogenic trace gas emissions from airborne measurements over Sacramento, California in spring 2009. Atmospheric Chemistry and Physics 11, 705-721.

We have included this reference.

Pg 20185 line 14. Why not use _CO2 and _CO rather than _x and _y?

It was pointed out to me (by the Editor) that the use of chemical formulae as physical quantity symbols (e.g. “CO2”) is not appropriate. IUPAC (International Union of Pure and Applied Chemistry) suggests the symbol y for gas mole fraction. For clarity, I use y for CO2 and x for CO.

Pg 20185 lines 19-20. Please reference the production of CO from VOCs.

We have added a reference for this in the revised manuscript (i.e. Granier et al., 2000.)

Pg 20185 lines 23-27. This is an awkward sentence.

We have changed the sentence in the revised manuscript.

Pg 20187 line 10. Please reference how the biodiesel/biogasoline content is known.

We have added the reference in the revised manuscript (IEA, 2014).

Pg 20187 line 11-12. Why is the vehicle biofuel emission ratio higher than for vehicle fossil fuel emissions? This is counterintuitive, since vehicle CO emissions have largely been reduced by the use of catalytic convertors, which one would expect to be similarly effective no matter the source of the fuel. It would be worth clarifying in this discussion what exactly is meant by biofuel. Does it refer only to biodiesel/biogasoline used in vehicles, or to open fires (e.g. for home heating), or other sources? The CO emission ratio can be expected to vary wildly across these different combustion types.

When we talk about biofuel CO2, we mean all combustions of solid (e.g. wood, waste, charcoal, municipal renewable waste, bagasse, vegetal waste and dung), liquid (e.g. biodiesel, bio gasoline and black liquor) or gaseous (from compost or cattle farm) bio-material. It does not include large scale biomass burning. We have included this information in the manuscript in the Introduction.

In Pg 20187 line 11-12 we mean the total biofuel contribution and not the biogasoline/biodiesel contribution. In our emission inventory, biofuel contributions have a larger emission ratio (CO/CO2) than traffic emissions. We clarify this in the revised manuscript.

Pg 20188 line 9 and throughout. It is common practice to use ppm and ppb rather than μmol/mol-1, etc. This could be explained at the first use if the journal prefers SI units.

We (as well as the journal) prefer to use SI units throughout the text.

Pg 20190, lines 1-2. Excluding NMHC oxidation to CO seems problematic. This is discussed to some degree later in the paper, but the potential problems with excluding this CO source should also be mentioned here.

In the revised manuscript, we mention the consequences here, but we keep the discussion in the last part of the paper.

Pg 20190 line 16 to end of section. This whole section is very heavy on equations, and it is difficult to follow. The equations that are given in the main text explain how each individual parameter is determined, but the equations that are used in the 6 different tracer combinations used in the analysis are found only in the appendix and table. There's no explanation of what these 6 different tracers are or why they were chosen in the main text. Table 2 helps a little, but requires the reader to recall what all the variables mean to interpret it. A paragraph that explains why these tracer combinations were chosen and what the assumptions and prior information required for each are would be helpful.

In the current manuscript we discuss the advantages and disadvantages of the different tracer methods in the Introduction. However, we agree that it would be helpful to shortly summarize the tracer-methods and underlying assumptions used. We therefore included additional comments in the revised manuscript to give the reader a better overview and clearer structure.

Pg 20191 line 7-8. Please explain why CH4 minimum values were chosen as background.

We have chosen to use CH4 minimum values as a proxy for atmospheric boundary layer mixing, since this parameter is frequently measured in parallel to CO2 and CO (e.g. in ICOS). CH4 sources in Germany are often distributed rather homogenously resembling the homogenous Radon flux (for Heidelberg see Hammer, 2005). This also implies that CH4 sinks do not vary on a diurnal time scale.
If however large point-sources (e.g. emissions related to fossil fuel extraction, processing, transportation or distribution) or large sinks of CH4 exist in the catchment of the station, it would be better to use Radon concentrations. We have tested the use of Radon and CH4

for our three model sites and have not found any relevant differences between the tracers used. This is not shown in the manuscript, but we included a respective comment in the revised manuscript.

Pg 20193 lines 8-9. Are these overestimates at the different sites results from the model study done here, or from previous work?

The principle sign of the misassignment of fuel CO2 using CO2 only is an outcome of our model study and also intuitive. The quantification of the over- and underestimation is a result of our model study.

Pg 20193 lines 18-20. Shouldn't this be the first paragraph of the following section?

This paragraph was thought as a short connecting passage between the paragraph 3.1 and 3.2, but we agree that it is better to have it as first paragraph of the following section We have changed this in the revised manuscript.

Pg 20194 line 12. "Until now"? Does this refer to the initial part of this study described above, or to previous research?

It refers to this chapter. We have changed this in the revised manuscript.

Pg 20195 lines 10-19. Indeed, the model shows that the δ13C method works quite well, but how reliable are the δ13C values of the sources (which are of course critical to the success of the method)? This is discussed in a later section, and that section should be referred to here.

We have referred to the later section in paragraph 3.2 in the revised manuscript.

Section 3.3.1. It is also worth noting that the CO2-only method bias varies seasonally.

We have noted this effect earlier (p. 20193, l.4-9). In section 3.3, we discuss only the actual sensitivity of the methods on certain parameters. We therefore do not note the seasonality at this subsection.

Section 3.3.3. Have the authors considered the impact of C3 vs C4 plants in the biospheric δ13C signal? In urban areas, lawn grass may be C4, which would have a large impact on the biospheric δ13C, and could introduce quite large biases to the δ13C method.

As shown in Table 1, we have included a variation of the isotopic signature of the biosphere. The assumed seasonal cycle of the isotopic signature stems from the relative proportion of C3 to C4 plants. The lawn grass in Europe is mainly C3, but plants such as corn or sugarcane, is C4. As is pointed out by the reviewer, misassignment of these isotopic biospheric values leads to a bias in the d13C(CO2)- based method. This can also be seen in Figure 4f and is discussed in Section 3.3.3.

Section 3.3.4. This section should reference the previous papers that have discussed the importance of Δ14C precision. Also, only a handful of 14C labs (AMS or gas counting) routinely achieve 2‰ precision, and most 14C labs report significantly poorer precisions. Finally, ±2‰ in Δ14C is closer to ±1 ppm than ±1.5 ppm in fuel CO2.
Please clarify what is meant by the last sentence of this paragraph.

We reference Turnbull et al., (2007) in the revised manuscript and now state that 2 ‰ precision is the best momentarily achieved precision rather than the typical precision. We also correct the conversion of permille D14C to fuel CO2. Finally, the last sentence describes that the 14C(CO2)-method is not able to detect biofuel CO2 and therefore leads to a bias in the total fuel CO2 estimate. We re-wrote this sentence in the revised manuscript.

Pg 20200 line 2-4. Since production of CO from VOCs is not included, this will be a systematic bias in one direction, not random.

This is correct if we obtain the ratio RF=CO/CO2 from bottom-up estimates. However, if we determine the ratio RF experimentally (by ^{14}C calibration), CO offset will be the sum of natural and anthropogenic CO offsets and only the variation of the natural CO ratio will lead to errors. These errors will be bi-directional and depend on the natural CO offset during the time of “calibration” and the natural CO offset at every moment. Nevertheless, we agree that in both cases (determining RF using bottom-up or top-down approaches), the bias due to non-fossil CO offset is not random. We now mention this in the revised manuscript. Note also that, e.g. the CO soil sink will bias the fuel CO_2 estimate in the other direction than CO sources and therefore the sign of the total bias depends on the net CO flux.

Pg 20200 line 8-12. As mentioned earlier, the 14C measurement method described in the McIntyre paper doesn’t appear to be applicable to continuous atmospheric 14CO2, so there’s no obvious reason for this uncertainty of 5‰ to be chosen.

We agree and refer to our comment above.

Pg 20203 lines 6-8. Please reference the use of afternoon-hour-only data in models.

We have added a reference.

Pg 20203 lines 15-17. Earlier in the paper much larger fuel CO2 values were given for the urban areas. Here I think the 1-2 ppm value indicates mid-afternoon values only, whereas earlier the values included nighttime? Please clarify here and in the earlier discussion.

This is correct. We clarify this in the revised manuscript.

Pg 20204 lines 7-12. Please reference the ICOS program.

We have already given the link to the ICOS homepage in the manuscript and the project number in the acknowledgements.

Pg 20204 lines 13-16. The authors may also want to refer to the Turnbull 2011 paper (reference given above) that uses aircraft grab samples.

We have added the reference.

Pg 20204 line 27. Please just state the number of samples used, rather than using n/24, etc. When only 1 monthly sample is used to determine RF, what is the uncertainty in this value, and how does this influence the results?

In the ICOS program we have only 24 samples available, but in other programs and settings we might have more samples available and we also discuss this later in Table 3. Therefore, we would like to keep it general. One calibration a month means two 14C measurements (background and sample). Taking this sample randomly would lead to a typical R_F misassignment of about ±2 ppb/ppm. This leads to large misassignments in the CO-based fuel CO2 estimate of about ±3 ppm.

Pg 20206 lines 3-10. When only a single background is used, all values could be biased if that background value is biased.

This is correct and the respective additional uncertainty is already included in the values given in Table 4. We will explicitly state this shortcoming in the revised manuscript.

Sections 5 and 6. The discussion and conclusions are thoughtful and interesting, but somewhat repetitive of each other and of the results section. They could be combined and/or substantially shortened.

We have shortened the discussion and conclusions in the revised manuscript.