Review of Wenger et al., Atmospheric radiocarbon measurements to quantify CO2 emissions in the UK from 2014 to 2015

This paper describes a set of radiocarbon (14C) measurements from two sites in Ireland and the UK, and uses these measurements to determine is also measured, and the potential and challenges of using CO as a ffCO2 tracer are considered. Transport model simulations are performed for all 3 C isotopes in CO2 and these model results are used to (a) diagnose the influence of nuclear industry 14C production and heterotrophic respiration fluxes on 14C and hence the calculated ffCO2 values for this region and (b) compare the modelled and calculated ffCO2 and CO values. The results demonstrate that although previous research has shown that nuclear industry 14C emissions can be problematic for ffCO2 studies in the UK, many parts of the UK are not overly influenced by this problem. recently added fossil fuel CO2 (ffCO2). Carbon monoxide (CO) from the same sites

I reviewed an earlier version of this paper. Thank you for the exceptional improvement over the previous version. This revision is clear and easy to follow. It is now apparent that the overall concept and strategy are strong, and the paper is appropriate for publication in ACP. I have a number of comments, but overall only minor revisions are needed before publication.

We would like to thank the reviewer for spotting many important mistakes, inaccuracies and misleading statements. We appreciate the large amount of effort and time that was obviously spent on this review as well as the review of the earlier version. We are glad that the reviewer found this version to be an improvement. This would not have been possible without the very fair and helpful comments of the reviwer and we are very grateful for it.

Specific comments:

Line 27. Core Writing Team reference – I think this is a reference to the most recent IPCC Working Group I document, but the reference is incomplete.

We thank the Reviewer for spotting this error, the reference has been corrected

”The level of carbon dioxide (CO2) in the atmosphere is rising because of anthropogenic emissions, leading to a change in climate (IPCC, 2014; Le Quéré et al., 2018).”

Lines 44-46. There are many more studies where 14CO2 has been used to estimate ffCO2, many of them in places other than just those listed, including in Asia.

We did not intend to imply that the references are a list of all ffCO2 estimate studies. We corrected the sentence to make clear that this list is only a subset of the 14CO2 studies that have been performed.

“Burning fossil fuels, therefore, causes a depletion in $^{14}CO_2$ that can be observed in the atmosphere, a phenomenon known as the Suess effect (Suess, 1955). Previously, $^{14}CO_2$ has been used to estimate CO_2 from fossil fuel burning ($ffCO_2$) in, among other places, the USA, Canada, New Zealand as well as some European countries (Bozhinova et al., 2016; Graven et al., 2012; Levin et al., 2003; Miller et al., 2012; Turnbull et al., 2009a; Vogel et al., 2013; Xueref-Remy et al., 2018)”

Line 63. “as each source emits with a different CO:ffCO2 ratio. Please reference this statement.

We have referenced this statement by adding a reference to the EMEP/EEA air pollutant emission inventory guidebook.

” However, using a CO_{enh}: $ffCO_2$ ratio to estimate higher frequency $ffCO_2$ can be challenging to implement even when using a well-calibrated ratio because the ratios of different sources and sinks impacting each measurement can vary considerably, as each source emits with its own CO : $ffCO_2$ ratio (Adams et al., 2016).”

Lines 78-81 and lines 91-95. So the flasks are collected from the 185 m height, but the in situ CO observations are from a lower height. It seems possible that the difference in CO mole fraction between flasks and in situ could simply be due to the difference in sampling height rather than a scale issue. Please comment on this. CO in the flasks, or are you confident that the offset is only a scale issue? If the former, then is there any possibility that other species from the same flasks could also have a problem? Please clarify.In general, I would have thought that it would be better to use the CO from the same flasks/height as the 14C measurement, since they are being used together. How would using the flask CO data (rather than the in situ data) change the results and interpretation?

This is a very valid concern that the Reviewer addresses It would have been ideal to have high frequency CO observations taken from the same height taken as the flask samples. Unfortunately, at the time of the study this was not the case. There are many other sites beside TAC that are equipped with both AGAGE instrumentation (CO on CSIROP scale) and NOAA flask sampling (NOAA scale) that show a similar pattern. At all these sites both types of samples are taken from the same inlet height, this makes us very confident that most of the difference is caused by the calibration scale differences. Theoretically, the calibration scale offset was linear during the study period and could be corrected for. However, we wanted to avoid this since it might be misunderstood to be an official conversion factor. There are time dependent drifts in the scale differences, using a correction factor from our work for other data sets could lead to large errors. Both NOAA and CSIRO are working on trying to resolve the calibration scale issues for CO.

At the Reviewer suggested we included a comparison of the time matched differences from the other gases. 1. The differences between the in-situ observations at 100m, and 185m for CH4 6.41ppb (0.33 %) and CO2 0.99ppm (0.24 %) and 2. the difference between the in-situ observations and the flask measurement both sampled at 185m for CH4 3.47ppb (0.18 %) and for CO2 0.42ppm (0.11 %). This shows that there is a certain amount of variability added by using a different height to compare it with flask measurements. However, the differences are nowhere near as large as the differences we can see between the flask and in-situ CO observations (up to 10%). Although, the CO measurement uncertainty is also a bit higher, so that probably explains part of the variability as well.

”Observations of CH_4 and CO_2 at the two heights were similar (less than 0.35% difference) within the same hour the flasks were taken indicating that it was acceptable to use the CO observations at 100m. A comparison of the concentration of CH_4 and CO_2 in the flask samples vs. the respective time matched in situ observations at 185m showed good agreement (less than 0.2% difference).”

Line 83. Miller et al 2012 used free tropospheric measurements from the same aircraft flights as background, not upwind sites. Be careful about calling sites “unpolluted”, as all sites will be influenced by local or regional pollution to some extent.

Line 85. I believe you mean Turnbull et al 2015, not 2014.

We thank the Reviewer for being very thorough and spotting these mistakes in the references. Both instances have been corrected.

” Different types of sites have been utilised as background in previous studies: relatively unpolluted sites upwind of significant fossil CO_2 sources (Lopez et al., 2013), high altitude observations (Bozhinova et al., 2014; Levin and Kromer, 1997), free troposphere observations from an aircraft (Miller et al., 2012; Turnbull et al., 2011) and a mildly polluted site upwind of the polluted site (Turnbull et al., 2015).”

Lines 134-135. Please edit for grammar.

The Authors edited the whole text again and hope this as well as other remaining mistakes have been corrected.

Lines 153-155. It is not strictly correct that fractionation discriminates 14C twice as much as 13C . Farnhi et al (2017) discuss this in detail and show that the approximation used in S&P 1977 is sufficient given the current uncertainties in the 14C measurement.

Fractionation effects discriminate against ^{14}C approximately twice as much as for ^{13}C (Fahrni et al., 2017; Stuiver and Polach, 1977).

This is another very good and valuable comment from the Reviewer, we edited the text to make clear that it is only approximately a factor of 2 and added the appropriate reference.

”Fractionation effects discriminate against ^{14}C approximately twice as much as for ^{13}C (Fahrni et al., 2017; Stuiver and Polach, 1977).”

Line 154. I think you mean ⍰14C, not ⍰14C!

Unfortunately, the formatting has been lost on the comment, we have tried to implement it to the best of our ability. We changed the ambiguous sentence:

The $\Delta^{14}C$ is normalized to a $\delta^{13}C$ value of -25 ‰,…

To :The $\delta^{14}C$ is normalized to a $\delta^{13}C$ value of -25 ‰ to obtain $\Delta^{14}C$,…

Normalising “should” remove reservoir specific differences? Are you suggesting that it might not be effective? If so, please explain.

We removed the “should” as it made the sentence imprecise.

Line 173. Please discuss a little more about how the background was calculated. Later you indicate the the background uncertainty for ⍰14C is ~4‰, which is much higher than the 14C measurement uncertainty which is quoted for the TAC site measurements. Were the MHD measurements done to lower precision, or is uncertainty calculated in some other way? Please add a figure (perhaps in the supplementary material) that shows the MHD 14C data and the median 14C values that were used as background.

It is important that the reader understands how to reproduce the calculation. To aid this the text has been lightly edited and at the suggestion of the Reviewer we added a figure showing the background observations vs. the rolling median value used in the calculation to the supplementary. In addition to this, a table including all the values used in the calculations in section 3.3 was added to the supplementary material as well, to aid transparency and understanding.

The 4‰ indicated later in the text is an estimate of how much the variability in Δ_{bg} adds to the total uncertainty of the $CO_{2\,ff}$ value. We estimated that this influence of the Δ_{bg} variability to the total uncertainty is twice as large as the average measurement uncertainty of the $\Delta^{14}CO_2$ observation in MHD (2‰).

” The rolling 15 percentile value (± 20 days) of the high frequency CO_2 observations at MHD (background site) was used as $CO_{2\ bg}$. For the Δ_{bg}, the rolling median value of the $\Delta^{14}CO_2$ flask measurements at MHD were calculated within a time window of ±20 days of the Δ_{obs}. Figure S.9 in a plot of the supplementary shows the MHD $\Delta^{14}CO_2$ observations and the rolling media value of the data used as Δ_{bg}.”

”All values used in the calculation of $CO_{2\ ff}$, including the Δ_{obs} and the Δ_{bg} and the correction terms have been included in the supplementary material in Table S10.”

Line 178. Suggest adding a brief note about why only nuclear and heterotrophic respiration are considered, and not other sources such as the ocean.

We added the following sentence to acknowledge that there are other corrections and why we decided to focus on the Biospheric and nuclear correction.

” In addition to these two correction terms explained below, other work (Graven et al., 2012; Turnbull et al., 2009b), investigated corrections for cosmogenic ^{14}C production and for the ocean atmosphere CO_2 exchange, for both corrections the modelled values are generally smaller than the uncertainty of the $\Delta^{14}CO_2$ measurements and they were therefore considered negligible for this work.”

Line 205 and line 217. Why is 10^15 ‰ first mentioned, then 7.3x 10^14 then used? And please explain where this value comes from.

We have changed the text, so the values are consistent, clearly identifiable and referenced. It was calculated using the Activity of pure 14C.

”($\Delta_{nuc} \approx 7.3x10^{14}$ ‰ (Bozhinova et al., 2014))”

Line 231. Please include (in the supplement) a table of the 14C values and sampling info (lab numbers, time/date sampled, etc)

This is a very good idea that we have tried to implement, we included the measurements as well as the corrections used in the $CO_{2\ ff}$ calculation for TAC in Tabel 10.

”All values used in the calculation of $CO_{2\ ff}$, including the Δ_{obs} and the Δ_{bg} and the correction terms have been included in the supplementary material in Table S10.”

Line 231. “Daily mean values” are these 24 hour means, or are they only daytime? If only part of the day, please indicate which hours of the day are used. What time of day were the flask samples taken, and does that match with the modelled time of day used?

Daily mean values (24h) are plotted, as it is very hard to take diurnal biospheric fluxes into consideration in the modelling set up we used. The back trajectories we used show an integration of where the air came from in the 30 days before the observation (according to the model). It assumes that emission fluxes do not change over those 30 days. Using emission data at a higher time resolution than these 30 days just does not result in very meaningful results. For CO_2, if an emission file with diurnal emissions is used for example on a summer day in the afternoon (high uptake of CO_2 by plants), this would result in a very low modelled CO_2 value in our set up because it assumes this high uptake took place for the whole of the 30 day period not just during daytime. One of our Authors Emily White has developed a system to use time weighted back trajectories in NAME but the method was not ready when we preformed our study.

E. White 2019: Quantifying the UK's carbon dioxide flux: an atmospheric inverse modelling approach using a regional measurement network, Atmos. Chem. Phys., 19, 4345–4365, https://doi.org/10.5194/acp-19-4345-2019, 2019.

The sampling times varied (listed in the table S10) but generally the site personal was instructed to sample after 12am so the boundary layer would be higher. However, there are some instances when sampling took place earlier in the day.

We added a clarification that it is 24h daily mean values "Daily mean values (24h)"

Lines 240-246. How does the uncertainty in 13C play out in terms of the 14C measurements that are the focus of this paper? Does it matter?

The uncertainty in the modelled values do not have an influence on the observations and how they were processed. However, uncertainty in the modelled 13C could make the modelling of the 14C more uncertain if it is directly propagated. It is important to measure the 13C value in the atmosphere when taking 14C observations so mistakes made in the ffCO2 calculations can be minimized by using the observed 13C value. However, it is certainly true that some assumptions are made in the calculation of ffCO2 that can cause additional uncertainty if you have a highly variable 13C value. The authors feel that since the error estimation for both the modelled values and the observed ffCO2 calculation is rather conservative, uncertainties caused by 13C variability should fall within the total estimated uncertainty.

Line 249. See previous comment about uncertainty in background 14C values.

Measurement uncertainty and estimated uncertainty in a model is not equivalent and often factors of the measurement uncertainty are chosen to be more representative of the true uncertainty of a system. This more conservative approach (a factor of two of the measurement uncertainty) was used in the error estimation here.

Line 266. Clarify that "ffCO2 equivalent" is the correction terms in equation 4.

We explained and defined the term fossil fuel equivalent in the text now, hopefully this makes the sentence easier to understand. While it is true that it describes the corrections in equation 4 in this sentence, it is more universally describing the relationship between 14C depletion and CO2 from fossil fuel burning.

" The term fossil fuel equivalent is used to describe how much recently emitted fossil fuel would have to be present in a sample to cause the equivalent depletion in ‰ in ^{14}C, the exact conversion from one to the other depends"

Lines 276 – 279. From figure 6, it appears that the nuclear correction is 100-200% larger than the ffCO2 value at the TAC site, yet the text implies (but doesn't explicitly say) that it is a much smaller correction. Please add some more detail about the relative contribution of the nuclear (and biosphere) corrections to the ffCO2 calculation, and the implications for the reliability of ffCO2 for this site and the UK.

The Reviewer might have misinterpreted figure 6. It shows the average ratio of how much nuclear emissions enhance the atmospheric 14C vs. how much the atmospheric 14C is depleted by fossil fuel emissions at a given location during a 1 year period. This means the plot shows good and bad locations to use the radiocarbon method for ffCO2 calculations. The figure only shows the ratio between the two, without any indication of how large either the ffCO2 or the 14C from nuclear emissions is. The nuclear corrections are indeed generally quite small, close to nuclear industry sites 14C emissions clearly dominate over ffCO2. The largest part of the UK, is in light to medium blue (5-1), this does indeed mean that 14C enhancement is dominant or equivalent to the depletion caused by fossil fuel burning. However, this is not because the 14C enhancement from nuclear sites is very large, but because the fossil fuel emissions are so small. The darker blue values indicating locations where fossil fuel emissions should cause a depletion in 14C observations that is on average larger than the nuclear correction that was applied. It is important to remember that figure 6 shows only average yearly value, not if a specific time was good for an observation or not. The main problem of the ffCO2 method in the UK is that the observed ffCO2 signal is not often large enough to give meaningful results.

It would be helpful to show the ffCO2 values calculated for each sample – this could be an additional panel in either figure 2 or figure 3 or a separate figure (it would be nice to show the time series of ffCO2 from the model as well). Currently the only place ffCO2 values are shown is in figure 4, which is useful but we can't tell which point relates to which correction in figure 3.

The Authors think that this is an excellent idea and have added the calculated ffCO2 both to figure S.8 as well as table S.10 in the appendix. Table S.10 contains details about the observations as well as the corrections applied.

Line 288. Is there a typo in the boiler inspection date? Seems like it should be 10th June 2014, not 10th July 2014 – the July date couldn't have caused a problem at TAC on June 13th!

This is corrected now, thank you for noticing.

" on the 13th June 2014"

Lines 295-300. Does the model include emissions from continental Europe, and if so, what is the quality of those emission estimates? I'm wondering if the problem is model transport during this low wind speed period, or if there is an issue with the emissions from this region as well?

The emission inventory used in the model does contain emissions for the whole modeling domain. The main problem is indeed the transport and the random mixing in the model. We added the following sentence to clarify why the model preforms badly during extended periods of low wind speed.

" in extended period of low wind speeds the modelled wind speed and direction have considerable uncertainty and variability due to the dominant influence of local terrain features that are sub-grid scale and therefore not resolved"

Line 305. Explain and/or reference the 1 ppm to 2.5 ‰ relationship.

We have added a better explanation of how the 14C observations correlate to the expression of fossil fuel equivalent.

" The term fossil fuel equivalent is used to describe how much recently emitted fossil fuel would have to be present in a sample to cause the equivalent depletion in ‰ in ^{14}C, the exact conversion from one to the other depends"

Lines 308-309. You say that the observed ffCO2 uncertainty is relatively large while the UK ffCO2 emissions are relatively low. I think you mean that the ffCO2 mole fractions (not emissions themselves) are relatively low – ie signal to noise is poor.

The Reviewer is absolutely correct, we have corrected the sentence.

” However, the uncertainties associated with the observed ffCO$_2$ are relatively large, while the ffCO$_2$ mole fractions observed at TAC are comparatively low. ”

Line 313-314. Please reference the previous work on CO:ffCO2 ratios and their variability.

We have added a sentence about previous work on the CO:ffCO2 ratio based on 14C observations.

” Other studies have found a wide variety of CO$_{enh}$ / ffCO$_2$ ratios, generally older studies have a higher CO$_{enh}$ / ffCO$_2$ ratio such as Turnbull et al., 2006 with 20 ±5 ppb ppm^{-1} or Vogel et al., 2010 with 14.8 ppb ppm^{-1}, whereas more recent studies in Europe have found similar CO$_{enh}$ / ffCO$_2$ such as Vardag et al., 2015 in Germany 5 ±3 ppb ppm^{-1} and Ammoura et al., 2016 in France 3.0-6.8 ppb ppm^{-1}. ”

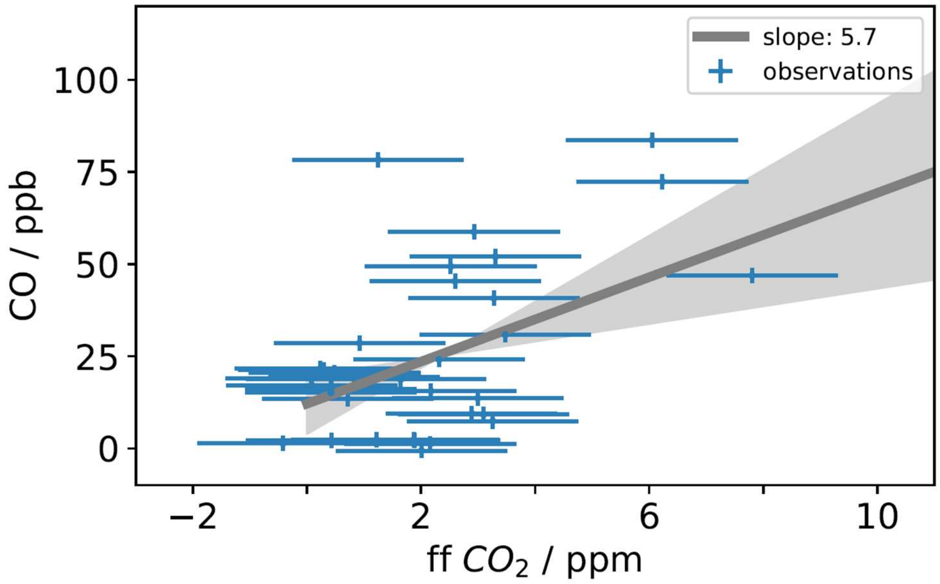
Line 315. Do diesel cars really have that low an emission ratio? And what about petrol cars? There are a number of recent studies that show onroad CO:ffCO2 ratios of around 5-15 ppb/ppm (depending on the country, emissions controls, etc).

Well these are the emission factors from 2014, before the Volvo emission fraud scandal, so the real emission factors might be a bit different now. During cold starts in cold weather diesel engines have lower CO emissions than petrol, however it all depends on running conditions, temperature and the vehicle size, we tried to rephrase the sentence to make the statement more clear.

”According to the NAEI 2014, UK gas power plants (1.0 ppb (CO) ppm (CO$_2$)$^{-1}$) and cars (0.5 ppb (CO) ppm (CO$_2$)$^{-1}$) under ideal driving conditions have low emission ratios, while larger vehicles preforming a cold start or accelerating on the motorway can have an emission factor an order of magnitude larger. ”

Lines 321-327. The fit in figure 5 looks to be strongly constrained by the two high points, and if they were excluded it looks like you would get a much higher ratio. Maybe those two points are samples where there was a strong local influence in the sample (e.g. a car idling nearby)? Further, curve fits to noisy data, and data with uncertainty in both axes is tricky. What kind of linear regression was used; were the ffCO2 uncertainties accounted for in the regression; how would excluding the two high points change the slope and it’s interpretation?

The two high points in blue the reviewer is referring to are the two very high values measured in November 2014. We excluded them from the data analysis as they would otherwise dominate the regression and because the NAME model was not able to capture them well due to extended periods of low windspeed over Europe. They were left in the figure 5 by accident, as a previous version of the paper included a CO ratio analysis both with and without those two points. The updated version of the plot no longer contains these two data points.



The linear regression uncertainty was estimated by the bootstrapping method presented in the paper. A normal linear regression without uncertainty estimate was used for the calculation specifically the Python, stats, linreg function. The linear regression was then recalculated 10000 times. Each time the dataset was allowed to vary (randomly) within the observation uncertainties then the resulting varied dataset was randomly resampled. The resulting spread of linear regression data was then used as an indication of the uncertainty of the method.

Lines 328-329. Reference this statement.

”Therefore, to maximise the scientific value of low frequency ffCO_2 observations, ffCO_2 has been used to calibrate the CO_{enh}/ ffCO_2 ratio for an individual sampling site ($CO_{enh} = CO_{obs}-CO_{bg}$) (Ammoura et al., 2016; Levin and Karstens, 2007; Miller et al., 2012; Turnbull et al., 2006; Vardag et al., 2015). ”

Lines 336-340. This argument that variability in the CO:ffCO2 ratio is due to variability in traffic CO:ffCO2 ratios needs some more justification. There are quite a few studies of onroad emission ratios that could be referred to. Those studies show that indeed, individual vehicles do vary considerably in their CO:ffCO2 emission ratio, but it is not clear that the variability in individual vehicles translates to variability in tower measurements where the traffic signal is a mix of many, many vehicles.

We added a statement to clarify why it is reasonable to expect a large variability in the CO even in an integrated signal. The paper states that while the CO ratio is similar to the model value over long time periods (spatial and temporal integration) this is not necessarily true at any individual point in time (no temporal integration and spatially more variability). Most studies that specialize on road transport emissions make observations in highway tunnels or next to highways over reasonable length of times. This means that they observe a specific type of traffic, in one place and then integrate this over time. As there is less spatial variability (where the air came from) it is not unreasonable that even shorter temporal integration (weeks, months) would lead to a good average value. A tall tower observation site will see much more variable emission sectors and the regions that the tower is sensitive to will change over time.

”While we expect to see an integrated emission signal from traffic at a tall tower site like TAC, each sample integrates air over a slightly different area with variable contributions from highways, country roads and city traffic. It is important to note that other source sectors have variable CO emission factors as well, for example in the sector domestic heat production, each individual boiler will have a different CO emission factor depending on the fuel source used and how optimised the operation conditions are. ”

Lines 340 – 344. If I understand correctly, the CO:ffCO2 ratio was derived from the 14CO2 measurements and CO measurements taken at the same time at the TAC site (figure 5). Then that CO:ffCO2 ratio is applied to the full time series of in situ CO measurements. The agreement between the model and observed ffCO2 is pretty decent (figure 4). So the results in figure S8 showing that the CO derived ffCO2 time series doesn't agree with the model seems at odds with everything else. This needs some more discussion and explanation. Some thoughts:

(a) the CO:ffCO2 ratio shown in figure 5 is skewed way too low by those two high points that might be locally influenced. If they were excluded, you'd get a higher ratio and therefore the CO derived ffCO2 time series would have a smaller magnitude and match the model better.

This was an error on our part. Figure 5 did erroneously contain the two high points from November 2014 even though they were not included in most of the CO ratio analysis. This has been updated. In Figure S.8 the ffCO2 calculated with the CO ratio used 5.7 ppm / ppb as the ratio which was calculated without the values in November. Table 1, still contains the result of all versions of the CO ratio analysis, it shows that including the two points in November would indeed lead to a higher ratio of 6.5 ppm / ppb.

Data	R^2	ppm / ppb
All	0.9 (0.5-0.9)	6.5 (4.8-7.9)
All (not Nov)	0.5 (0.2-0.7)	5.7 (2.4-8.9)
Winter only	1.0 (0.7-1.0)	6.6 (4.6-8.0)
Winter only (not Nov)	0.7 (0.1-1.0)	4.7 (1.0-10.1)

(b) Does figure S8 show just the daytime when (presumably) flasks were sampled, or does it include nighttime data? I'd be surprised if the model does a good job at night, so if nighttime data is shown it could be confusing things.

Figure S8 shows 24 daily mean values. Including nighttime data. The reviewers concern about this is valid. Generally the model preforms better if the boundary layer height is larger (during the day), additionally the sampling was skewed to daytime only sampling meaning it might not be representative for daily mean values. We aim to investigate this in the next measurement campaign and have added a disclaimer to the text mentioning this potential bias.

”In addition to that $\Delta^{14}CO_2$ observations at TAC have predominantly been timed to take place in the afternoon, this might bias the calculated CO ratio to be more representative for daytime observations. ”

Liens 348-350. As in a previous comment, figure 6 seems to show a very high contribution from the nuclear industry, whereas the text asserts that the influence is small. This needs to be clarified!

See comment above for Lines 276 – 279.

Line 358. In this paper, the flask samples were very short grab samples, and Turnbull et el 2012 used a 1 hour integrated sample. So where does the 3 hour integration come from? What's the justification for choosing 3 hours vs 1 hour or some other time period? This paper doesn't make any comparison of integrated samples vs grab samples, so I don't see how the outcomes of THIS study can lead to a recommendation that integrated samples are better – although it is probably true that integrated samples have advantages in many situations.

Integrated samples would be better represented in the model. This means that corrections applied for the nuclear influence and the Biospheric disequilibrium for the ffCO2 calculation would be more representative for the individual measurement. A 1h integration is much better than a grab sample, a 3h integration period would remove the observation from being affiliated with the closest back trajectory in time, multiple back trajectories could be aggregated, potentially smoothing out model errors. We agree that this is all very theoretical but believe it is justified to make the recommendation since it is clear that improving the application of the corrections also improves the reliability of the ffCO2 method.

Lines 360-362. What is the justification for suggesting conditional sampling? As in the previous comment, it's probably useful to do that, but THIS study doesn't add any new information about the usefulness of doing so. It's worth noting that conditional sampling might be a great idea, but it is also likely to be much more difficult to do!

Our study shows that while on average, many parts of the UK are not ideal for sampling 14C for ffCO2 calculations (Figure 6, and generally how low the observed ffCO2 is), individual samples might still give significant results. It therefore follows that being able to time the sampling to catch favorable sampling conditions more often would be beneficial, especially for such an expensive measurement.

Lines 364-369. This should first be presented in the methods and results, not just in the discussion section.

We added the description of the forward modelling to the method section.

”To simulate the concentration of a substance in the modelling domain, theoretical particles are released at the emission source location (point sources and area sources) with a rate that is relative to the emission source strength. ”

Figure 6 only shows the ratio of nuclear correction to ffCO2. Whereas the magnitude of ffCO2 will also be important in considering site locations, so that the signal is large enough to be measurable.

Lines 370-374. This study doesn't seem to provide any evidence that the method would work better at the city scale. Further, if discussing London, please indicate London's location in figure 6.

This is correct. However, the average nuclear correction is fairly uniform outside of the immediate vicinity of a nuclear industry site, this is why we are confident in stating that the main problem is the low average ffCO2 and sites closer to high emission regions would result in more frequently significant ffCO2 observations. Doing city scale emissions means sampling closer to a high emitting region, which results in a larger signal. We have added the location of London to figure 6.

Comment on grammar and language. I noted several very minor typos and grammar issues as I read. I have not commented on them all since they are likely to change in revision, but I suggest a good check through for such issues before resubmission.

Comment on authorship. Have the current authors considered whether the NOAA and INSTAAR scientists who contributed the 14C and other flask measurements should be included as authors on this paper?

Yes we have asked both NOAA and INSTAAR about how they would prefer their contribution to this work to be recognized.

The manuscript "Atmospheric radiocarbon measurements to quantify CO2 emissions in the UK from 2014 to 2015" by Wenger et al. makes interesting use of measurements of atmospheric 14C in CO2 in order to attempt to estimate fossil fuel emissions from the United Kingdom. This is an interesting and potentially useful approach and the publication of the data and the model comparisons would certainly be beneficial. The methodology appears to be thorough and robust. Whilst it is disappointing that the measurement uncertainly appears to prohibit a thorough understanding of the emissions, the work carried out merits documentation. Whilst, previously, much of the manuscript text was unclear and made it difficult to assess the method and results, the text has been significantly changed since the original submission. The authors have clearly made every effort to improve the manuscript based on the comments provided in the previous reviews, and the study is therefore much easier to follow than before.

The manuscript now reads well, with only a few technical corrections remaining. The figures are generally clear and well chosen and the methods and models used within the manuscript are appropriate for such a study. The terminology is consistent and the chosen equations are clear and appropriate.

I recommend publication of this manuscript subject to the following minor and technical changes.

The Authors would like to thank the Reviewer for the nice comments and the helpful corrections. We would especially like to express our gratitude for spotting that there was an error in Figure 4. We are very grateful to have been able to correct the mistake and are still in disbelieve that none of us spotted this.

Minor changes & technical corrections:

Page 1, line 12-13: should be “as emissions from fossil fuels, which do not contain 14CO2, cause a depletion…”

Page 1, line 14: radiocarbon-derived fossil fuel CO2 (ffCO2)

Page 1, line 20: COenhanced has not been defined in the abstract. Better to describe fully e.g. ‘by deriving a constant ratio of CO enhancements to ffCO2 for the mix of…’

Page 2, line 39: ‘to disentangle’ or ‘of disentangling’

We have implemented all the changes exactly as stated.

Page 2, lines 64-70: These sentences need rewriting a little. Make it clear that this paragraph describes the forthcoming sections. The second sentence in particular is unclear: ‘In this study we use these observations to…’

We have changed the last paragraph and hope it is now clearer that it describes the forthcoming sections.

”As part of the Greenhouse gAs Uk and Global Emissions (GAUGE) network (Palmer et al., 2018), weekly $^{14}CO_2$ measurements have been made at two sites between July 2014 and November 2015: Tacolneston, Norfolk (TAC, 52.51°N, 1.13°E), a site that is influenced by anthropogenic sources in England and Mace Head, Ireland (MHD, 53.32°N, -9.90°E), a background site. In this work, we present a way to model the isotopic composition at TAC and MHD and compare the modelled data to the observations. The $^{14}CO_2$ measurements are then used to calculate $ffCO_2$ at TAC. The need for this radiocarbon-based calculation of the $ffCO_2$ to be corrected for the influence of $^{14}CO_2$ from nuclear power plants and the biospheric disequilibrium is also discussed. As an attempt to improve the temporal resolution of the $ffCO_2$ we define the CO_{enh}: $ffCO_2$ ratios at TAC and explore the potential for calculating $ffCO_2$ from high frequency CO observations. ”

Page 4, line 115: delete one instance of ‘CO2’, and comma after ‘trajectory’

Page 4, line 126: delete ‘was’

Page 5, line 133: ‘tool to investigate’

Page 5, line 135: The ‘CO2 obs’ in brackets here seems like it might be wrong. Should be ‘CO2 i’?

Page 5, lines 143 and 147: enhancement, or mole fraction?

Page 7, line 190: ‘assimilated’

Page 9, line 260: where -> were

Page 10, line 295: break up this sentence

We have implemented all the changes exactly as stated.

Page 10, line 302-303: Justify/explain this statement. The model would not

respond well to these conditions for what reason?

We have added a sentence to explain why the model does not preform well in long periods of low wind speeds.

" (in extended period of low wind speeds the modelled wind speed and direction have considerable uncertainty and variability due to the dominant influence of local terrain features that are sub-grid scale and therefore not resolved). "

page 10, line 304: replace 'modelled emissions' here. They're simulated

mixing ratios derived from modelling using reported inventories.

" In Figure 4 we present the $ffCO_2$ calculated with the radiocarbon method (Equation 4) from $\Delta^{14}CO_2$ observations at TAC station ($ffCO_{2\ observed}$) and compare it with simulated mixing ratios derived from modelling using emission inventories as described in Section 3.1 ($ffCO_{2\ simulated}$). "

Figure 4: I'm unclear about the corrections shown in this plot. If I understand

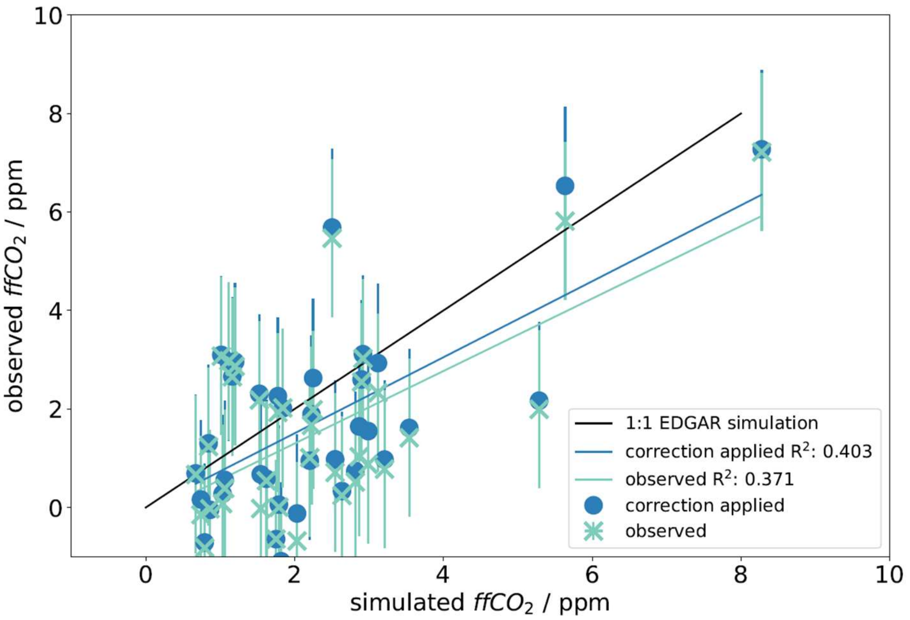
correctly they should match the corrections shown by the black dots in Figure

3, but that doesn't appear to be the case. Whilst all of the corrections is

Figure 3 are negative, the corrections in Figure 4 occur in both directions, and

never seem to be as large as those shown in Figure 3. Is this correct?

The Authors would like to thank the reviewer for spotting this flaw. Not one of us has noticed it in all the iterations. We could trace the discrepancy back to a mistake in the code, it was in an older file where the correction was calculated with a fixed value for the isotopic signature of the heterotropic respiration $\Delta_{hr,}$ instead of using a variable. This meant that when the value for the heterotropic correction changed, only one instance was corrected manually to the new value while the other instance was overlooked. We were able to find and correct the mistake and have added a new plot with the correct data points. In addition to this we have added a table to the supplementary data (S10), listing all the observational data as well as the corrections applied. This was done so the calculation would be more transparent and easier to reproduce.



Page 10, line 311: Remove question mark in this title

Page 12, line 349: 'impact on'

We have implemented all the above changes exactly as stated.

Page 12, line 365: only the nuclear 14CO2 signal is modelled, correct?

In this instance the ffCO2 at TAC was also simulated using Edgar 2010 and the back trajectories. I hope the addition in the text makes this more clear.

"If we take the average CO_{enh} / $ffCO_2$ ratio in TAC (5.7 ppb ppm^{-1}) as calculated above and multiply it with the high frequency CO_{enh} (as defined above), we get back a high frequency $ffCO_2$ time series for TAC. This time series of CO ratio derived $ffCO_2$ at TAC results in $ffCO_2$ values that are significantly larger than what the modelled $ffCO_2$ values suggest (simulated according to section 3.2, with the EDGAR 2010 fossil fuel emission map, Supplementary material S.8). "

Page 12, lines 367-369: clarify the sentence beginning 'These two

simulations are combined…' We have implemented this comment exactly as stated.

Page 12, line 374: You mean 'cost-intensive'? Yes we mean cost-intensive.