Review of Hu et al, ACPD, 2020. "Anthropogenic and natural controls on atmospheric δ13C-CO2 variations in the Yangtze River Delta: Insights from a carbon isotope modeling framework"

General Comments

As the authors point out, this is the first time (that I'm aware of) that CO2 and δ13C have been both modeled and measured for an urban area. I'm glad that they've attempted to tackle this issue, because applying multiple data streams (in this case CO2 and δ13C) within an atmospheric modeling framework should help us better understand urban CO2 sources and sinks. Although the inverse modeling methodology used is somewhat simplistic, I think it's a good start that can later be made more sophisticated. The forward modeling skill is only 'moderate' in my opinion (R^2 < 0.2 for CO2 and marginally better for δ13C), but this is perhaps unsurprising given the high noise urban environment and the relative proximity of the sampling site to local sources. Beyond the simulations, the authors carry out numerous interesting analyses using a combination of model results and observations. The topic they are addressing is important, the study is ambitious in scope, and is completely appropriate for ACP. There is no doubt that this paper represents a great deal of hard work in terms of both measurements, modeling, and analysis. All this said, I have some significant concerns and questions that need to be addressed before the paper can be published. I will outline my concerns immediately below and then provide detailed line by line questions and comments.

We thank the reviewer for the positive comments and detailed suggestions. We have made extensive revisions based on these comments.

1. δ13C data

My biggest concern is with the δ13C measurements presented. In particular, I am having a hard time understanding δ13C values that are greater (more positive) than -6 per mil, at times (Fig. 6), and more generally, summer afternoon values that appear to be close to -7 per mil (Fig. 7h). Looking at well-established background sites in the Northern Hemisphere such as Mauna Loa (from https://scrippsco2.ucsd.edu/data/atmospheric_co2/mlo.html) and La Jolla, δ13C values for July 2015 are around -8.4 and -8.3 per mil, respectively. δ13C from the NOAA network for Dec. 2014 at Mauna Loa (the last month available on their website) is -8.4 vs. -8.6 per mil for Dec. 2014 from the Scripps measurements, strongly suggesting that there are not significant offsets in the Scripps data. Given a rough starting point of ~ -8.4 per mil, it's very difficult to understand how a heavily polluted urban region where, as the authors say, the biosphere is a relatively minor component of fluxes, could raise a broadly representative background value of -8.4 to something close to -7 per mil. In principle, this could occur only with a large removal of CO2 from the atmosphere by net photosynthesis (leaving the atmosphere more enriched). To put a rough number on this, C3 plants fractionate approximately at a ratio of -0.05 per mil/ppm. This means that the CO2 levels seen at this study's measurement site would need to be roughly 28 ppm lower than La Jolla (for example) in July. While it's hard to tell exactly what the July 2015 daytime CO2 is, from Figures 3a and 7g it appears to be right around 400 ppm. For comparison, La Jolla CO2 for the same month is 397 ppm (the CarbonTracker backgrounds mentioned in the paper are very similar to this

value). Another possible explanation for high δ13C would be a source of CO2 with an isotopic signature heavier than the atmospheric value of ~ -8 per mil. Cement production (δ13C ~ 0), which is discussed extensively in the text, is of course such a source. However, even with the relatively large fraction of anthropogenic emissions as cement in the study region, the flux-weighted mean isotopic signature of anthropogenic emissions will still be much lower than the atmospheric values (~ -24 per mil, Fig. 12a). One could also ask the question if air from higher in the atmosphere might have higher δ13C levels (and be more appropriate to consider as a reference than La Jolla, e.g.). To answer that, we can look at data from 4 km asl from collected aboard aircraft (the NOAA site CAR at around 40 deg. N; see https://www.esrl.noaa.gov/gmd/dv/iadv/), where in July 2015, the δ13C is around -8.2 per mil, not much different than either La Jolla or Mauna Loa. The NOAA site LEF (a continental forested site with little industrial/urban influence) did record a δ13C of ~ -7.4 per mil in July, 2015. However, this isotopic enrichment was associated with a CO2 level of ~370 ppm, much lower than the hemispheric mean background. The only time prior to this paper that I've seen such enriched δ13C values have been in ice core samples (e.g. Francey et al., Tellus, 1999).

So the question is, why is this happening? I am not an expert in optical δ13C measurements, but reading the referenced paper Xu et al, ACP, 2017 as well as Ghasemifard et al, 2019 (Atmosphere) and Ghasemifard et al, 2019 (Aerosol and Air Quality Research), it is clear that the Picarro instrument used in this study requires significant corrections due to, among other things, water vapor. The fact that the very high values of δ13C are seen mainly in summer, while in winter (e.g. Fig. S4) the values seem much more reasonable, makes me wonder if the water corrections (which will be much more significant in summer) are playing a role here. It is worth noting here that even with frequent calibrations with reference air of well-assigned CO2 and δ13C, the fact that the reference gas is bone dry while the sample air is moist will pose a problem. I'm not saying here that water vapor is the explanation for the unreasonably enriched values during summer but rather suggesting this as a candidate for investigation. Almost more important than the fact that the data appears to be biased, I'm worried that there may be a seasonally varying bias in the data. Because the seasonality of the signal is an important part of the analysis in the study, it's important to make sure that, at the very least, any biases/offsets in the data are constant. I don't know why the data are as enriched as they are, but it's the authors' responsibility to convincingly explain the high δ13C values they observe.

We thank the reviewer for the very thorough comments and insights regarding the $\delta^{13}C$ signal. Here we discuss the main reasons for the relatively high $\delta^{13}C$ values and provide a rational for our revised approach.

1. $\boldsymbol{\delta}^{13}C\text{-}CO_2$ background observations

Recently, Ghasemifard et al. (2019) showed that hourly $\boldsymbol{\delta}^{13}C\text{-}CO_2$ values at Mount Zugspitze, the highest (2650 m) mountain in Germany, varied between -7‰ and -12‰ in the winter for 2013. During two especially clean air events (in October and February) at Mount Zugspitze, the $\boldsymbol{\delta}^{13}C\text{-}CO_2$ was approximately -7‰, during which the CO_2 mixing ratios varied between 390 and 395 ppm. This is consistent with our estimates using strategies 2 and 3. We have

added it on lines 497-501.

2. Noise in the empirical estimate of the background values

In Figure 6, the hourly δ^{13}C-CO_2 background values were derived by combining hourly δ^{13}C-CO_2 observations with the Miller-Tans approach. This derived result is subject to large fluctuations at the hourly time scale because in equation 4, the derived δ^{13}C-CO_2 background will have similar hourly variations with the atmospheric δ^{13}C-CO_2 observations. For example, the derived hourly δ^{13}C-CO_2 background values fluctuated by more than 2‰ within a single day (shown in Figure 6). These large fluctuations are physically unrealistic given our understanding of background values observed at remote sites. For these reasons, we used the smoothing and fitting technique to provide a best estimate of the slow varying background component.

3. The effects of water vapor on the $\boldsymbol{\delta}^{13}$C-CO_2 IRIS measurements

As described in our previous work (Xu et al., 2017), we found that the Picarro IRIS $\boldsymbol{\delta}^{13}$C-CO_2 measurement has some dependence on the ambient water vapor mixing ratios. This bias was quantified based on sensitivity analyses of increasing water vapor mixing ratios while measuring an air cylinder with a constant $\boldsymbol{\delta}^{13}$C-CO_2 value. We observed an increasing trend of 0.46‰ of measured $\boldsymbol{\delta}^{13}$C-CO_2 for a 1% increase in water vapor when its mixing ratios exceeded a value of 2.03%. This dependence is shown below in Figure R1.

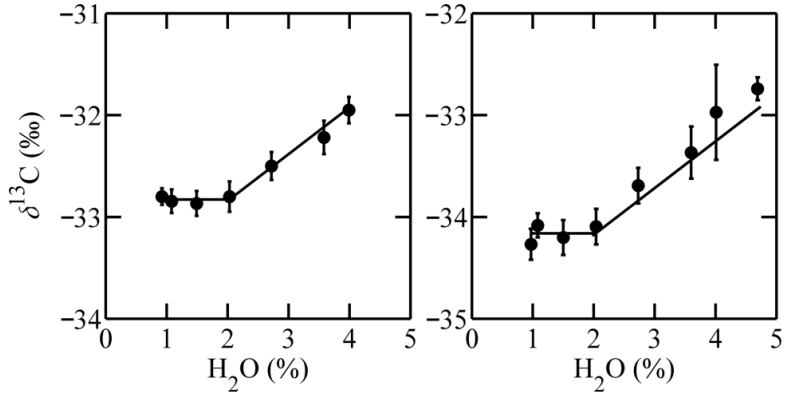

Figure R1. Dependence of the observed δ^{13}C on the H_2O mole fraction. The lines represent error bars are ±1 SD of 1 min averages. The data in the left panel were obtained on 1 October 2014 using a 439 μmol mol^{-1} standard gas cylinder with a true δ^{13}C value of −32.8 ‰. The right-hand panel is for 10 June 2015 using a 488 μmol mol^{-1} standard gas and the true δ^{13}C value of −34.1 ‰.

The correction procedure is as follows:

$$\delta^{13}\mathrm{C} = \delta^{13}C_{true},\ \mathrm{C}(H_2\mathrm{O}) \le 2.03\%,$$
$$\delta^{13}\mathrm{C} = \delta^{13}C_{true} + 0.46‰,\ (\mathrm{C}(H_2\mathrm{O})\% - 2.03\%),$$
$$\mathrm{C}(H_2\mathrm{O}) > 2.03\%,$$

where the $\delta^{13}C_{true}$ is the true isotope delta value, δ^{13}C is the measured isotope delta value (after a two-point calibration), and C(H_2O) is water vapor mole fraction. This sensitivity test indicates that the true δ^{13}C value is not sensitive to water vapor when water vapor mole fraction is lower than 2.03% (Figure R1, R2). The δ^{13}C value is biased high when water vapor mole fraction is higher than 2.03%, which should be corrected following this calibration

procedure. Based on above equations and our observed values during this study we find that the $\delta^{13}C$ values will not be subject to this calibration in the winter and the highest hourly corrections were 0.74‰ as observed during the summer when water vapor mole fraction is higher (Figure R2b).

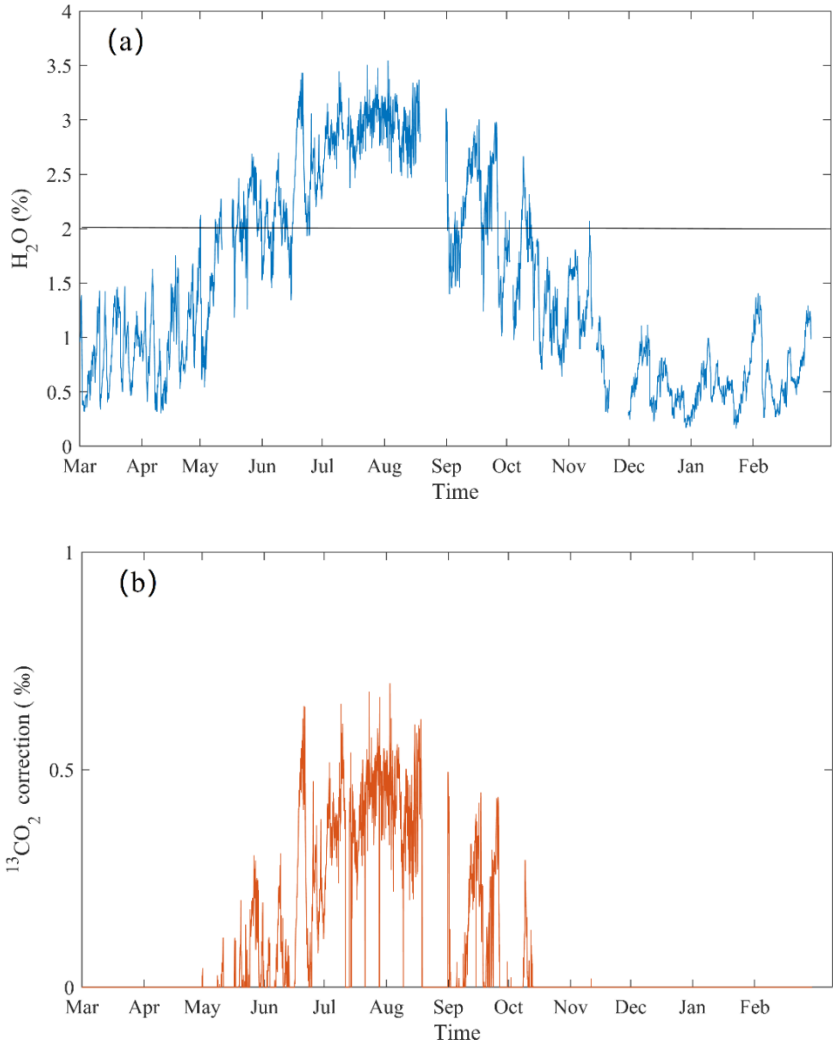

Figure R2. (a) Hourly variations of observed H_2O mixing ratio, and (b) corrections on hourly $\boldsymbol{\delta}^{13}C\text{-}CO_2$ observations.

We added the following description "We note that the $\boldsymbol{\delta}^{13}C\text{-}CO_2$ IRIS (model G1101-i) measurements are sensitive to water vapor concentration. Sensitivity tests reveal that the $\boldsymbol{\delta}^{13}C\text{-}CO_2$ IRIS measurements are biased high when water vapor mole fraction exceeds 2%. The data presented here have been corrected following the procedures outlined in Xu et al. (2017)." on lines 153-156 for clarification.

4. Other biophysical factors

We agree that three other potential reasons can help to explain the enriched (or high) $\boldsymbol{\delta}^{13}C\text{-}CO_2$ values as discussed on lines 201-210 and 485-488 including: (1) vertical gradients of $\boldsymbol{\delta}^{13}C\text{-}CO_2$; (2) fractionation associated with ecosystem photosynthesis; and (3) enrichment associated with the CO_2 derived from cement production.

5. Our best estimate of the background $\boldsymbol{\delta}^{13}$C-CO_2 values

Although the derived hourly δ^{13}C-CO_2 background values fluctuated by more than 2‰ within a single day, here we also calculated the daily minimum background $\boldsymbol{\delta}^{13}$C-CO_2 in winter, which is displayed with yellow line in Figure R3, the average was -8.15‰ and comparable with WLG winter observations of -8.55‰. Based on the above analyses and discussion, we believe that our best estimate of the $\boldsymbol{\delta}^{13}$C-CO_2 background values in the WRF-STILT model framework are derived from smooth curve fitting.

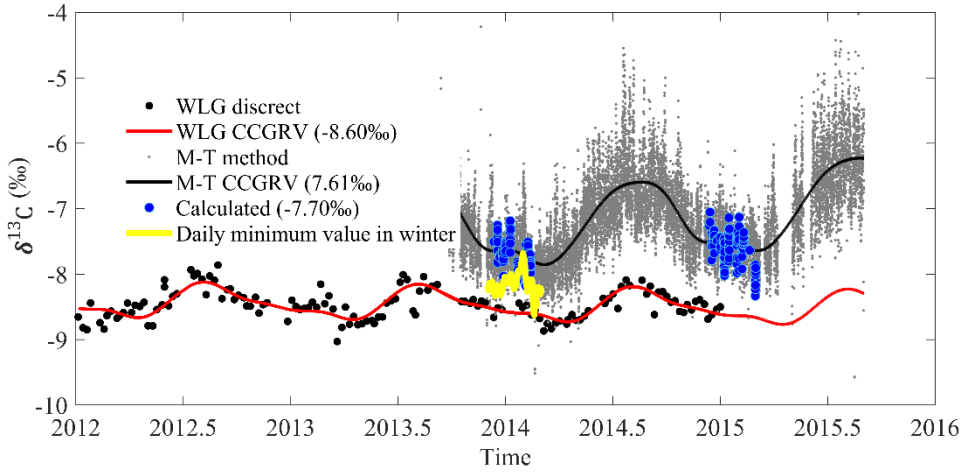

Figure R3. Comparisons among three strategies for calculating the background $\boldsymbol{\delta}^{13}$C-CO_2, with figure caption the sane as Figure 6 in the main text.

2. *Daytime analysis*

The vast majority of both forward and inverse model analyses have focused on afternoon data. The main reason for this is that atmospheric transport models generally have a much harder time simulating shallow nighttime boundary layers often with strong vertical gradients (where fewer model levels are available to capture vertical gradients) than they do simulating higher mid-day boundary layers, when the PBL tends to be well-mixed. Another big advantage of focusing on mid-day data is that the daytime turbulence in the PBL serves to integrate fluxes over a much larger upwind region (in time and space). Thus, conclusions about sources and sinks, especially when using data from just a single site, are much more likely to be spatially representative. I would like to see the model-data comparisons and other analyses using afternoon selection criteria (12-16 hr, e.g.). Even if model-data comparison statistics do not radically improve, other analyses, such as the enhancement proportions of different sectors could change by minimizing the influence of very local sources. With model-data comparisons, in particular, it could be that the model performs similarly for the full record as it does for just the daytime part. If so, this would be an interesting finding.

Done as suggested. Here we chose to display the results of daytime (10:00-16:00, local time) to represent the periods with higher mid-day boundary layers. We compared the time series of CO_2 mixing ratios, enhancement proportions of different sectors, scatter plots, and monthly scaling factors as described below.

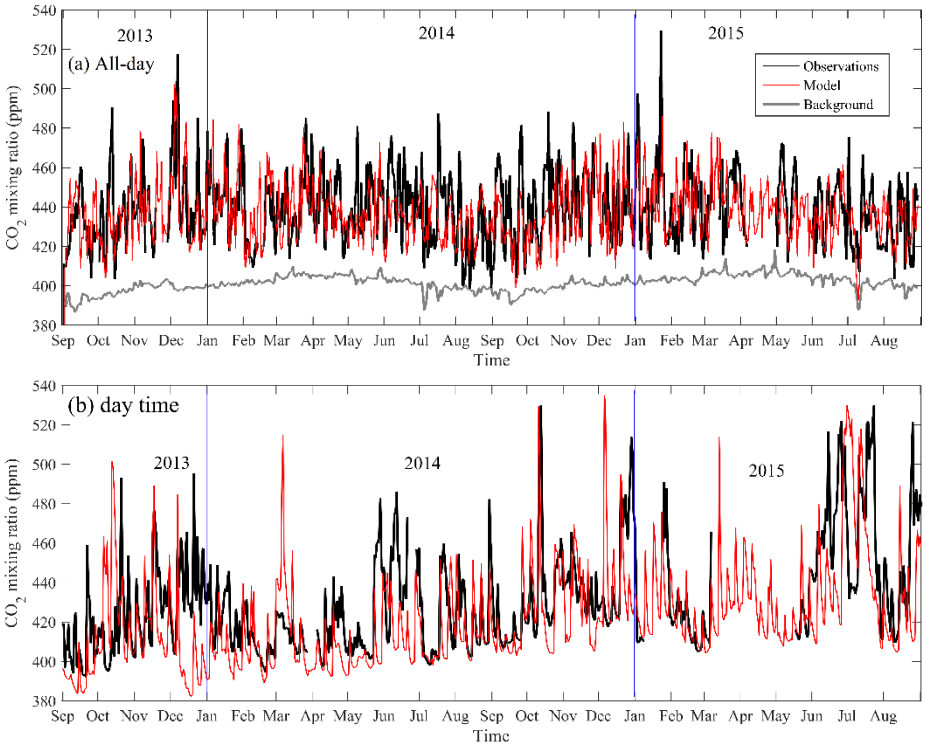

Figure R4. Comparisons of CO_2 daily averages for (a) all-day and (b) daytime.

We calculated the RMSE, R and Mean Bias (MB) for daily averages for all-day and daytime only, which were 18.68 ppm, 0.44 and 2.68 ppm for all day averages, and 25.21 ppm, 0.38 and 10.74 ppm for day time.

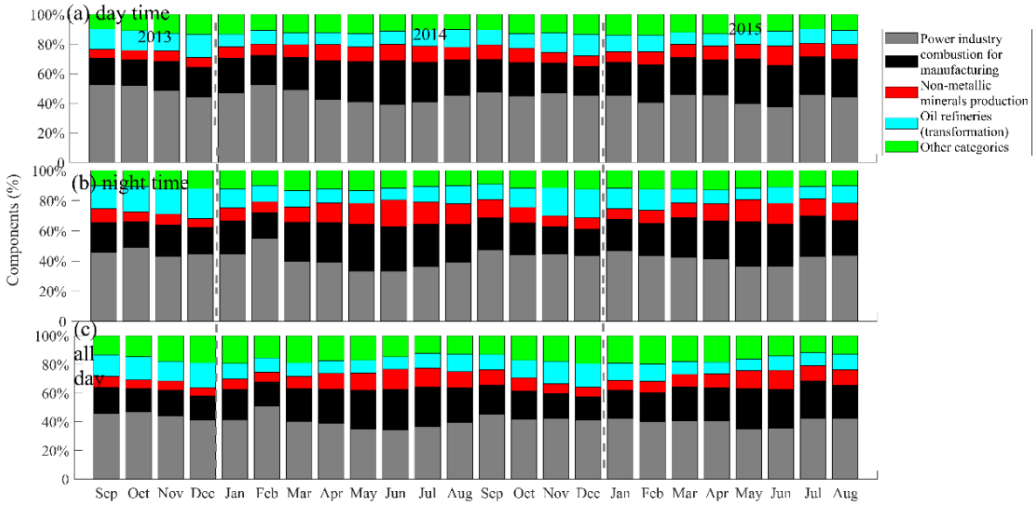

Figure R5. Components of anthropogenic sources for (a) daytime (10:00-16:00), (b) nighttime (22:00-06:00), and (c) all-day (0:00-24:00).

We also re-examine the comparisons of atmospheric CO_2 components for daytime (10:00-16:00) and the nighttime (22:00-06:00) results. Figure R5 indicated that they have similar trends, but with different magnitude. We attribute this to differences in the source footprint for day vs nighttime.

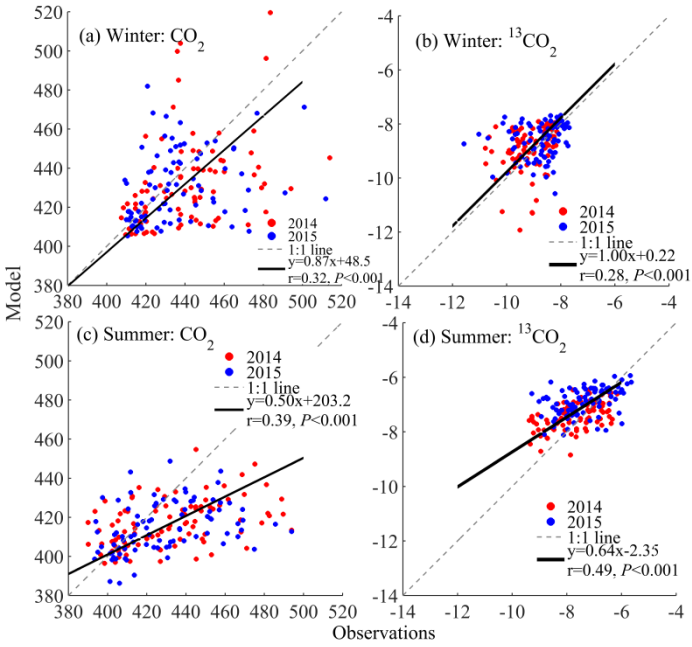

Figure R6. Scatter plots of observed versus modeled (a) winter time CO_2 mixing ratios, (b) winter time $\boldsymbol{\delta}^{13}C\text{-}CO_2$, (c) summer time CO_2, and (d) summer time $\boldsymbol{\delta}^{13}C\text{-}CO_2$ for both years, here these dots are day-time (10:00-16:00) averages.

We also performed comparisons by only choosing the daytime observations. The results indicated that daytime CO_2 mixing ratio simulations in the summer were slightly underestimated. This caused $\boldsymbol{\delta}^{13}C\text{-}CO_2$ to be overestimated (Figure R6). The simulations for winter generally captured the trends for both CO_2 and $\boldsymbol{\delta}^{13}C\text{-}CO_2$ when the biological CO_2 enhancement played a relatively small role compared to anthropogenic emissions. The larger bias in the summer could result from the relatively coarse spatial-temporal resolution (aggregation error) of the Carbon Tracker biological CO_2 flux, which was 1×1 degree with three-hour average. As shown in Figure S3, the spatial distribution of land use is far more heterogeneous. This will smooth the stronger biological CO_2 signals by averaging it over the large 1×1 degree grid, while the urban biological CO_2 flux occurs at much finer spatial scales and likely varies at shorter time intervals. We add this description on lines 529-542.

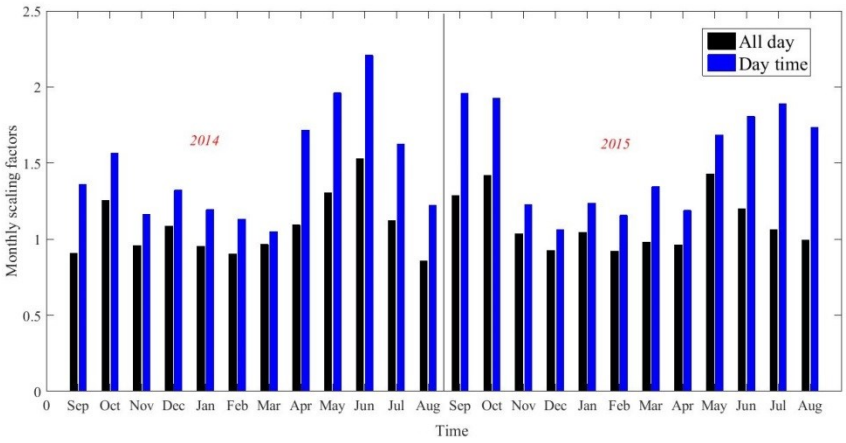

Figure R7. Derived monthly scaling factors for all day and only day time.

The monthly scaling factors derived by only using daytime CO_2 observations are displayed in Figure R7. We should note that they are pretty close to results with using all-day CO_2 observations from October to March, when the biological flux is smaller compared to the main growing season. During April to September some large inconsistencies are evident. The negative daytime NEE will cause the "observed anthropogenic CO_2 enhancement" in equation 8 to be biased high and result in larger monthly scaling factors.

We added the descriptions for clarification on lines 454-464 as "The *a posteriori* results indicate that the annual scaling factors were 1.03 ± 0.10 for 2014 and 1.06 ± 0.09 for 2015. The monthly scaling factors derived from using daytime and all-day observations are also shown in Figure S4. These factors vary seasonally with higher values observed in summer. When using daytime values only, the scaling factors were much larger than the all-day values. This can be seen in Figure 3 by comparing the simulated and observed CO_2 mixing ratios. We should note here that the larger scaling factors based on the daytime data could be caused by bias in the *a priori* daily scaling factors used to generate the hourly CO_2 emissions (Hu et al., 2018b); the monthly anthropogenic averages; and bias in negative biological CO_2 enhancement. Since our study is mainly focused on the seasonality of all-day observations, the monthly scaling factors derived from the all-day approach will be used for the following analyses."

3. Equations
As detailed below, the details of the equations are hard to follow, especially in terms of what is simulated and what is measured. I have made some suggestions to clarify the notation.
Thank you very much for these suggestions. We have adopted many of these changes. These changes are mainly on lines 175-259. We added subscripts of "obs" and "sim" to represent observations and simulations, respectively.

Specific Comments
L46 change v432 to v4.3.2
Done as suggested.

L47-48. "and constrained the anthropogenic CO2 emission categories." This is misleading. The scaling factor approach only constrained the total anthropogenic emissions. The isotopic data were used constrain the cement fraction to some extent.
Thank you for catching this. We have revised this as "constrained the anthropogenic CO_2 emission categories" with "constrained the total anthropogenic CO_2 emission"

L50. "performed well" This is debatable and subjective. The R^2 values for fits to CO2 data were less than 0.2. If you want to comment on WRF performance, please quantify instead of saying "well".
We changed "performed well in reproducing" with "can generally reproduce".

L54. Delta_s has not been defined at this point, so you need to say what it means.
We added "(the mixture of $\boldsymbol{\delta}^{13}$C-CO_2 from all regional end-members)" following Delta_s to

define it.

L58. Change 'plants' to 'plant'
Done as suggested.

L78. Change 'by' to 'from'
Done as suggested.

L85. I don't think this is the correct IPCC reference. What you want to cite here are the IPCC guidelines on emissions calculations, not the IPCC report on the science of climate change.
Done as suggested. We changed it to "IPCC (Intergovernmental Panel on Climate Change): 2019 Refinement to the 2006 IPCC Guidelines for National Greenhouse Gas Inventories, available at: https://www.ipcc-nggip.iges.or.jp/ public/2019rf/ (last access: 24 April 2021), 2019"

L88. Change 'into the inversion of global biological CO2 flux' to 'into the estimation of biological fluxes in atmospheric inversions.' It's not the biological fluxes that are being inverted. Also global is not appropriate here because you are talking about high fossil uncertainties at local scales.
Done as suggested. We revised it as "These large uncertainties are propagated into the estimation of biological fluxes in atmospheric inversions".

L122. Change 'have recently be' to 'have recently been'
Done as suggested.

L123. Change 'inversion has been' to 'inversions have been'
Done as suggested.

L129. Change 'power' to 'the power'
Done as suggested.

L151. The NOAA/ESRL lab you refer to should now be referred to as NOAA/GML (NOAA Global Monitoring Laboratory).
Done as suggested.

L176. A) does "ms" in CO2_ms refer to measured? Or is the left-hand-side of eq. 1 just a simulated quantity. This seems to be the case from the text above, but to make that clearer, I suggest adopting more intuitive notation such as 'CO2_sim'. B) Delta_CO2 on the right hand side of eq. 1 might be better written as Sum{i=i,n}[Delta_CO2]i to be consistent with eq. 2.
Here ms refers to simulation. Thank you for pointing this out. Changed as suggested. We have revised the expressions from equation 1 to equation 6.

L177. Change 'hands' to 'hand sides'
Done as suggested.

L179. Again, I think the notation should be clarified. The left hand side of eq. 2, as I understand it, is the simulated value of atmospheric δ13C. Make that notation consistent with that for the simulated value of atmospheric CO2, e.g. δ13C_sim.

Done as suggested.

L183. As mentioned above, in eq. 3, instead of Delta_CO2 use Sum{i=i,n}[Delta_CO2]i. Then it becomes very clear what the definition of delta_s is: the enhancement-weighted mean isotopic value of all sources/sinks.

Done as suggested.

L184. Change 'the mixture' to 'the enhancement-weighted mean', which is more precise.

Done as suggested.

L194-195. I don't agree with: "background air masses should originate from the free atmosphere at heights of 1000 m or higher above the ground". In general there shouldn't be a specific altitude requirement for background air. A more general definition when conducting regional studies would be that background is the concentration or isotope ratio of air when the air enters the regional study domain, which is often determined using back trajectories. The back trajectories (often an ensemble as in the case of STILT) will exit the domain at a variety of altitudes, including possibly below 1000 m. Also, with regard to WLG in particular, this is a remote high altitude site that would be expected to be sampling free tropospheric air most of the time. While it is true that WLG is significantly to the west of the domain 1 border, given the size of the CO2 enhancements (and δ13C depletions) at the observation site, I would expect WLG to be a reasonable background site. As mentioned above, I think a more likely reason why WLG doesn't appear to be a good background site (more so in summer) has less to do with WLG and more to do with potential bias in the dataset. One quick experiment you can do is compare the CO2 values you extracted from CarbonTracker with the CarbonTracker values for WLG. I doubt there will be a substantial difference. If true, this would suggest that WLG should also be a reasonable background for δ13C.

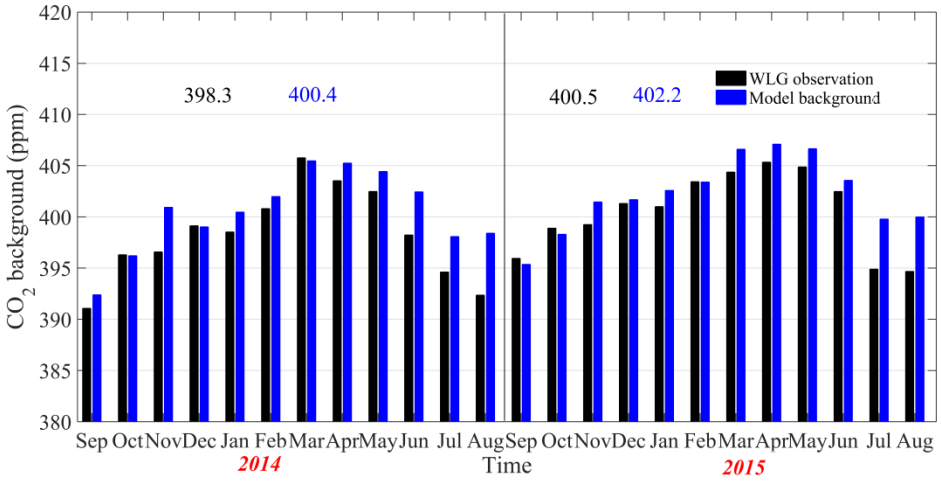

Figure R7. Comparisons of CO_2 background between WLG observation site and Carbon Tracker based model background.

Figure R8. View of WLG site.

Thank you for this comment. We agree that the back trajectories can also originate from below 1000 m, and is illustrated by footprint (Figure 2b, which is defined as the time-weighted values of released back trajectory particles below half of PBLH within each grid cells). Most of the back trajectories originated from the free atmosphere or higher above the ground. For clarification, we deleted the typo of "at height of 1000 m" and added "For example, based on the previous simulation results for the CO2 background sources, most of the back trajectories originate from the free atmosphere or 1000 m higher above the ground (Hu et al., 2019). Further, the footprint at the north/west edge of Domain 1 is relatively small, indicating that most back trajectories were observed above the planetary boundary layer height (hereafter PBLH)." on lines 202-206.

We also compared the CO_2 background between WLG site and the Carbon Tracker based model results. This comparison suggests there is ~2 ppm bias in the annual averages. The bias can increase to >5 ppm in the growing season (June-August). One possible reason for this is that the WLG site is dominated by grasslands. The large bias for summer indicates that the WLG site may not be a reasonable choice to define that background CO_2 value in summer.

L198-199. "can cause a high bias in the $\boldsymbol{\delta}$*13C-CO2 background when using this approach." Apologies if I am misinterpreting something, but using WLG is much lower (more negative) than the other background approaches. Why 'high bias'?*

Thanks for pointing it out, we revised the typo by changing "high" with "low".

L200. 'second approach'. It seems that the second approach is a very limited approach whose main purpose is to validate the 'third approach'. Thus, I would move this after the third approach and maybe not call it an 'approach' but say 'in order to test the second approach, we…'.

The "second approach" applied simulated $\boldsymbol{\delta}_{s\_sim}$ and CO_2 enhancement, while the "third approach" only used observed $\boldsymbol{\delta}_{s\_obs}$ and CO_2 enhancement, which makes them different.

L203. Here δ13C_a is referred to as an observed quantity, whereas in eq. 2 the same notation was used to refer a simulated quantity.

Yes, Here $\delta^{13}C_a$ is referred to as an observed $^{13}C\text{-}CO_2$, which was used to derive $^{13}C\text{-}CO_2$ background, we added "Note here that that $\boldsymbol{\delta}^{13}C_a$ represents the observed $\boldsymbol{\delta}^{13}C\text{-}CO_2$ not the

simulated $\boldsymbol{\delta}^{13}C\text{-}CO_2$ ($\boldsymbol{\delta}^{13}C_{a\_sim}$) as shown in equation 2." for clarification.

L205-206. 'minimize simulated CO2 enhancement errors' Are you referring to errors coming from NEE here? Is that why you chose the wintertime? If so, state this more explicitly.
The reason to choose wintertime is to minimize the influence from NEE including the photosynthesis and respiration. The reason to only choose bottom 5% wintertime CO_2 observations is to minimize both influence the from ecosystem NEE and anthropogenic CO_2 emissions. We revised this sentence as "We defined clean conditions as the bottom 5% wintertime CO_2 observations to minimize simulated CO_2 enhancement errors from both biological and anthropogenic CO_2 simulations on $\boldsymbol{\delta}^{13}C\text{-}CO_2$ background calculation" for clarification.

L209. Change 'equations' to 'equation'
Done as suggested.

*L209. Is Delta_CO2 here derived from CO2_obs – CO2_bg, or is it simulated? The first usage of Delta_CO2 in eq. 1 implies that Delta_CO2 is a simulated quantity, because you write about eq. 1: "CO2 was simulated as the sum of background (CO2_bg) and the contribution from all regional sources/sinks (**ΔΔ**CO2)". Perhaps eq. 1 is meant to describe that observed CO2 can be decomposed into a background component and the contribution from all sources. (Thus implying that Delta_CO2 is not simulated using footprints, but rather can be derived from observations and the background estimate.) However, on line 210 you write that in the third approach you do not need to simulated Delta_CO2_i. Please clarify.*
Here the Delta_CO_2 is calculated by using CO_2 observation minus CO_2 background. As described above, we have revised equation 1 and corresponding description.

"Note that $\boldsymbol{\Delta}CO_2$ is the sum of all simulated sources/sinks $[\boldsymbol{\Delta}CO_{2\_sim}]_i$ and represents the total simulated CO_2 enhancement. We use $\boldsymbol{\Delta}CO_{2\_obs}$ as the observed CO_2 total enhancement, which can be calculated by using the CO_2 observation minus the CO_2 background values." on lines 180-182.

L213. 'Similar methods…' I wouldn't say the studies referred to use similar approaches. For one, they don't involve isotopes. Second, in most of the approaches referenced, back-trajectories and information from remote sites were combined to determine background. In contrast, and very importantly, the δ13C background determined using method 3 is not independent of the observations themselves. This is a an important point because this is what allows you to define a background that fits so closely to the upper envelope of the observations, despite the fact that the smooth curve fit through the background (Fig. 6) is close to -6 per mil in the summer of 2015. As mentioned earlier, such values are not physically reasonable.
We agree that these references do not involve isotopes, and they generally applied observations to derive corresponding background for different trace gases. As explained above, In Figure 6, the hourly $\delta^{13}C$ background values were derived by combining hourly $\delta^{13}C$ observations with the Miller-Tans approach. This derived result is subject to large

fluctuations at the hourly time scale because in equation 4, the derived $\delta^{13}C$ background will have similar hourly variations with the atmospheric CO_2 observations. For example, the derived hourly $\delta^{13}C$ background values fluctuated by more than 2‰ within a single day (shown in Figure 6). These large fluctuations are physically unrealistic given our understanding of background values observed at remote sites. For these reasons, we used the smoothing and fitting technique to provide a best estimate of the slow varying background component.

L221-222. '1000 m above ground' Why wouldn't you use the concentrations from CarbonTracker at the altitudes where the back-trajectories exited the domain, instead of 1000 m agl? This may be a mis-interpretation or mis-reading of the Hu et al. 2019 methodology. As mentioned above, the background concentration should be taken from the lats, lons, and alts at which the ensemble of trajectories exit the domain. These values (500 in the case of STILT?) can then be averaged (and some aspect of their variance, perhaps the std. error of the mean, used to compute an uncertainty.)

Thanks for pointing out this typo. In the model framework, we choose the latitude and longitude to locate the backward trajectories, and as was found in previous studies most of these trajectories are at altitudes above 1000 m, we deleted "We used the averaged concentration at latitude and longitude when the released particles enter study domain 1" for clarification.

L223. 'hourly footprint function' say how these were calculated, or provide a reference here. Also, does this imply that footprints and back-trajectories were calculated for every hour of the data record? If so, state this.

Yes, the variable $\boldsymbol{\Delta}CO_{2\_sim}$ was derived by multiplying the simulated hourly footprint function with the hourly CO_2 fluxes (Hu et al., 2018a; b). Considering the diurnal variations of both anthropogenic and biological CO_2 fluxes, 168 footprints were obtained for each simulated hour. This accounted for the back trajectory of particle movement for 168 hours (i.e. 24 hours per day for 7 days) of transport. The 168 footprint will be multiplied by corresponding hourly CO_2 flux. We have added these descriptions on lines 240-243.

We revised equation 5 for clarification as

$$\Delta CO_2 = \sum_{i=1}^{168} flux_i \times footprint_i$$

We also revised the description on lines 254-257 as "where $flux_i$ (units: mol m^{-2} s^{-1}) corresponds to each CO_2 flux category simulated for each domain for a specific hour *i*, and footprint (units: ppm m^2 s/μmol) is the model simulated sensitivity of observed CO_2 enhancement to flux changes in each pixel. The *i* contains the hourly footprint during trajectory of particle movement for 168 hours as described above."

L233. When convolving the fluxes with the footprints to produce Delta_CO2 (here it's usage is clearly as a simulated quantity!), were hourly footprints for a single measurement convolved with hourly NEE to account for covariances in the diurnal patterns of both the footprints and the NEE? In other words, was there just a single footprint per measurement, summed over the

7 days, or were there 7x24 footprints saved? When focusing on simulations of biospheric CO2 this factor is very important. I suspect in your case, where anthropogenic fluxes without a significant diurnal cycle are dominant, neglecting this covariance between NEE and transport is reasonable. However, neither eq. 5 nor the text contain any of this information.

Yes, as answered in the previous comment, there were 7×24=168 footprints saved, which considered the diurnal variations of both biological and anthropogenic CO_2 flux. The biological CO_2 flux came from Carbon Tracker biological CO_2 flux, the anthropogenic CO_2 was derived with hourly scaling factors with EDGAR inventories. This method was the same as described in our previous paper (Hu et al., 2018b) in the reference list.

Hu, C., Griffis, T. J., Lee, X., Millet, D. B., Chen, Z., Baker, J. M., and Xiao, K.: Top-Down constraints on anthropogenic CO_2 emissions within an agricultural-urban landscape. Journal of Geophysical Research: Atmospheres, 123(9), 4674–4694, https://doi.org/10.1029/2017JD027881, 2018b.

L251. Here, you say that you used CarbonTracker values above 1000 m, which is different than what you said on line 221, where you implied that you said you used values at 1000 m. Please clarify. However, I'll repeat that there is nothing special about 1000 m. You should use the CO2 value from CarbonTracker at whatever altitude the trajectories left the domain.

We have revised this typo by changing "on 1000 m" with "above 1000 m". As described above, we tracked the latitudes and longitudes of particles when they entered domain 1. The reason we used the heights above 1000 m, is based on our previous study (Hu et al., 2018a) where we found that most of the released particles are above 1000 m height when entering domain 1.

L254. Say immediately which version of EDGAR you used.

Done as suggested, we added v4.3.2 after EDGAR.

L258. Modify 'the most up to date global inventory' to 'the most up to date global inventory with sectoral detail'.

Done as suggested, and thanks for pointing it out.

L264. It is not quite true that EDGAR v4.3.2 is only available for 2010. This is only true for the version with monthly resolution. Please add this qualifier.

Done as suggested, and thanks for pointing it out.

L266. Please explain more about how the factor of 1.145 was calculated. For example, was it based just on CT or did it use EDGAR for 2010 and CarbonTracker for 2014 and 2015? The former would be much better for consistency.

The scaling factor 1.145 was derived by dividing the anthropogenic CO_2 in CT in 2014-2015 by in 2010. We revised this sentence as "This scaling factor is based on Carbon Tracker, dividing the same anthropogenic CO_2 emissions for YRD in years 2014-2015 by that in 2010." on lines 291-292.

L269. Change 'posteriori' to 'a posteriori'.
Done as suggested.

L275. Again, with notation, please distinguish simulated (or bottom-up) delta_s as in eq. 6 with delta_s determined by a Keeling Plot, e.g.

$$\sum_{i=1}^{n} \delta_i \times p_i = \delta_{s\_sim}$$

We revised this as, "where δ_i is the $\boldsymbol{\delta}^{13}$C-CO_2 value from source category *i*, and p_i is the corresponding enhancement proportion (i.e. proportions of a specific enhancement i to total CO_2 enhancement). We define $\boldsymbol{\delta}_{s\_sim}$ as the simulated carbon isotope ratio of all sources to differentiate it from the observed $\boldsymbol{\delta}_{s\_obs}$".

L280. How much of non-metallic mineral production is cement? 60%, 90%, 99%? If nearly all, then I recommend stating this here and saying that from here on you will just refer to it as cement.
Here for the lack of detailed information, we simply attribute 100% of non-metallic mineral production to cement, and added such description on line 308-309.

L308. Change 'it' to 'them'
Done as suggested.

L310. At least for Hu et al., 2019, the inverse modeling approach was very different than that described here. It is a bit misleading to cite it.
Done as suggested, we deleted Hu et al., 2019.

L313. Again, confusing notation. Earlier, CO2_obs, was referred to as [CO2]. The "ms" subscript is still a mystery to me. In eq. 1. CO2_ms was total CO2 including all sources (and background). Here, Delta_CO2_ms is only the simulated anthropogenic enhancement. Why not make the subscript more intuitive, such as 'Delta_CO2_anth'?
Done as suggested, we change $\boldsymbol{\Delta}CO_{2\_ms}$ to $\boldsymbol{\Delta}CO_{2\_anthro}$.

L313. Are these terms monthly means? Or is a scaling factor calculated at high time frequency and then averaged to a monthly mean SF? Sargent et al., e.g., calculated the mean of all afternoon modeled and observed enhancements. Also, subsetting to afternoon data only, when modeled PBL heights are likely most reliable, is a necessary test.
Yes, these terms are monthly means. As described above, the reason to choose all-day averages is that both biological and anthropogenic CO_2 flux have strong diurnal variations (i.e. much higher in daytime and lower in nighttime for anthropogenic emissions), so if only use daytime observations, the derived scaling factors will reflect bias in both the *a priori* diurnal scaling factors and monthly averages of CO_2 emissions, so even the scaling factors is larger than 1, it does not only indicate the anthropogenic CO_2 is underestimated, it can also be caused by underestimation of diurnal scaling factors in daytime not the daily averages.

L329. As mentioned in the general comments, evaluating model performance with only afternoon data is strongly advised.
Done as suggested, as answered above, we evaluated model performance in daytime, which seems bias larger than all day performance.

L331. As mentioned above, I don't agree that R^2 < 0.2 equals "good" performance.
We revised "The model also performed well in simulating the monthly and seasonal variations of CO_2 mixing ratios" as "The model also captured the monthly and seasonal variations of CO_2 mixing ratios".

L348. I'm confused as to how you can compare 2014 and 2015, if 2014 was a full year but 2015 was a partial year. Are you only comparing the months that are common to both years?
Here as explained on line 157-160, "Further, for an annual comparison, we examined the period from September 2013 to August 2014 (Year 2014) versus September 2014 to August 2015 (Year 2015).", although 2014 and 2015 does not cover the whole 12 months in calendar year 2014 and 2015, they contain whole 12 months spanning two years for further comparison.

L349. Here you say that the emissions are the same for both years. This should be mentioned explicitly in the methods section.
Done as suggested, we added "This scaling factor is based on Carbon Tracker, dividing the same anthropogenic CO_2 emissions for YRD in years 2014-2015 by that in 2010." in methods section on line 292.

L353. I don't think you can conclude that meteorological differences between years were relatively small based on comparison of annual values. More analysis would be needed. For example, there could be a lot of seasonal cancellation that still results in similar annual averages.
Thanks for pointing this out. We revised it as "indicates a relatively small meteorological effect for the annual averages," for clarification.

L355. In terms of comparing NEE year to year, I think much more analysis is needed for this to be meaningful. First, NEE changes significantly on a seasonal basis. Second, because you are presumably looking at the full record, nighttime respiration is over-represented. Daytime values on the other hand incorporate both daytime NEE (GPP and ER) and respiration from the previous night (and probably more day/night cycles depending on the size of the domain and the windspeeds).
The hourly CO_2 enhancement contributed by NEE is also displayed in Figure 7b, which illustrated obvious diurnal and seasonal variations on lines 510-517.

L362. Note that the GPP product used is derived from the SIF product used. They are not really that independent. This can be seen in the shape of their seasonal cycles.
We agree that the GPP is derived from SIF observation and they are not independent.
We added "We note that GPP was derived from SIF, and as a result they share a similar

seasonal cycle." on lines 403-404 for clarification.

L370. While you may be able to ignore GPP during the winter, it doesn't mean you can ignore respiration. There will still likely be some NEE.
Yes, we agree that respiration in winter cannot be ignored as displayed in figure 7b. Here in this sentence, we only refer to the comparisons of photosynthesis between growing seasons and non-growing seasons.

L383. Change Washington to Washington D.C. (to distinguish from Washington state.)
Done as suggested.

L385. Change 'Eastern' to 'the eastern'
Done as suggested.

L389. "indicating greater…' Most likely, yes, but also trapping of emissions in the PBL will play a role. You cannot immediately transfer enhancements to emissions. So you need to qualify this by saying something like 'assuming similar windspeeds and PBL heights…'
We deleted the emission comparisons and revised this sentence as "Our enhancements were significantly higher than all of these previous reports of other urban areas"

L392. Say briefly how the percentages were calculated. I assume by convolution of each regions emissions with the footprints.
First, we calculated the CO_2 enhancement from each province by multiplying emissions in each provincial administrative boundary with corresponding footprint, and then divided the calculated CO_2 enhancement by total CO_2 enhancement for all area, we added "The CO_2 enhancement from each of the 5 zones were simulated by multiplying CO_2 emissions in each province with the corresponding footprint." on lines 257-259 in the Method Section for clarification.

L401. Change 'oil refinery' to 'oil refineries'
Done as suggested.

L412. How were the annual scaling factors calculated? As the unweighted (or weighted?) mean of the monthly ones? Also, the MSFs are never presented. What is their seasonality? This seems like a major result of the CO2 part of your analysis (and the 'inversion'). The monthly results should be discussed and/or presented.
Thank you for this suggestion. The annual scaling factors are the weighted mean of monthly values on lines 453-454 "We constrained the monthly anthropogenic CO_2 emissions by using the MSF method (equation 8) and computed the 12-month average to represent the years of 2014 and 2015.". As shown above in Figure R7, the scaling factors when using all-day generally varied around 1, while when only using daytime, they are generally larger than 1. Both results show seasonality. Both of them show obvious seasonality, which is lower from November to March, and higher from April to July and from September to October. We have added this figure in supplemental file and also added "We constrained the monthly

anthropogenic CO_2 emissions by using the MSF method (equation 8) and computed the 12-month average to represent the years of 2014 and 2015. The *a posteriori* results indicate that the annual scaling factors were 1.03 ± 0.10 for 2014 and 1.06 ± 0.09 for 2015. The monthly scaling factors derived from using daytime and all-day observations are also shown in Figure S4. These factors vary seasonally with higher values observed in summer. When using daytime values only, the scaling factors were much larger than the all-day values. This can be seen in Figure 3 by comparing the simulated and observed CO_2 mixing ratios. We should note here that the larger scaling factors based on the daytime data could be caused by bias in the *a priori* daily scaling factors used to generate the hourly CO_2 emissions (Hu et al., 2018b, Figure R9); the monthly anthropogenic averages; and bias in negative biological CO_2 enhancement. Since our study is mainly focused on the seasonality of all-day observations, the monthly scaling factors derived from the all-day approach will be used for the following analyses." on lines 453-468.

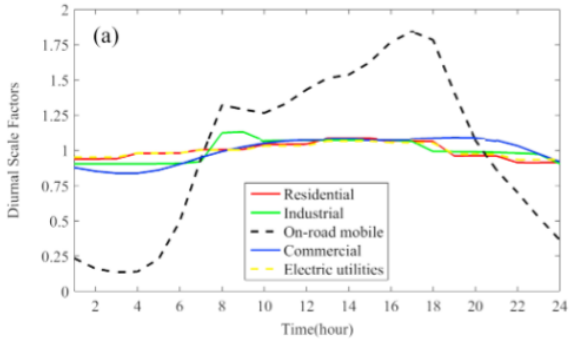

Figure R9. Diurnal scaling factors in Hu et al. (2018b), Figure 2a.

L415. Change posteriori to 'a posteriori' and 'for YRD area' to 'for the YRD area'.
Done as suggested.

L416. The last sentence on cement seems unnecessary and out of place, even if it is true.
We deleted the last sentence of "As noted, cement CO_2 emissions in the YRD is the largest regional source for global cement production". In 2013, the cement production from China accounted 60% of global total production, the second ranked country is India which accounted for 9% (USGS, 2014), and the YRD area accounted for 20% of national cement production in China, which is 20%*60%=12% in the global production (Xu et al., 2017; Yang et al., 2017). We cited the reference from USGS on line 132.
USGS (U. S. Geological Survey), 2014. Mineral Commodity Summaries 2013.http://minerals.usgs.gov/minerals/pubs/commodity/cement/.

L429. Using 'discrimination' to describe the isotope ratio of cement production is not appropriate, because it is not a fractionating process like photosynthesis is, e.g. You could instead say 'the isotopic signature associated with cement production…', e.g.
Done as suggested, we changed "discrimination" with "Enrichment of the isotopic signature".

L431. Change 'plants' to 'plant'
Done as suggested.

L432. Change 'than observed' to 'than the observed'
Done as suggested.

L438. Regarding Chen et al., 2006, the vertical gradients in that study are based on models. Observations were generally only at 20 m agl. In at least one example (Fig. 1) The vertical gradient looks to just about 0.1 per mil at mid-day. Moreover, at least for the summer months, the simulated values in the surface layer are more enriched due to photosynthetic drawdown.
Yes, the model work by Chen et al., 2006 seems $^{13}CO_2$ has small vertical gradient in the daytime because of the well mix of boundary layer, and large gradients of >1‰ was found between 20 m and 500 m in different years, which indicate the distinct signals in boundary layer and free atmosphere.

L439. Ghasemifard is not in the reference list.
Thanks for pointing it out, we have added it in the reference list.

L440. Saying that Zugspitze values were around -7 per mil for winter 2013 is not an accurate characterization of the Ghasemifard results. Looking closely at Fig. 2 of Ghasemifard et al. 2019, the average δ13C for DJF of 2012/2013 is at least 1 per mil lower than -7.0 per mil (even excluding pollution events). This extraordinarily high value of ~ -7.0 is only reached in two cases, once in October and once in January. Also note that -7.0 per mil is an unrealistically high value for Oct. 2012 in relation to similar high altitude data like that from the Scripps Mauna Loa record.
We revised the sentence as "Recently, Ghasemifard et al. (2019) showed that hourly $\boldsymbol{\delta}^{13}C\text{-}CO_2$ values at Mount Zugspitze, the highest (2650 m) mountain in Germany, varied between -7‰ and -12‰ in the winter for 2013. During two especially clean air events (in October and February) at Mount Zugspitze, the $\boldsymbol{\delta}^{13}C\text{-}CO_2$ was approximately -7‰, during which the CO_2 mixing ratios varied between 390 and 395 ppm.".

*L442. Saying that a clean air event pulls the δ13C down from -7 to -7.5 does not make sense. Pollution events (δ13C of fossil fuels average around -30 per mil) will make δ13C more negative. Here you are saying a *clean* air event makes the δ13C more negative.*
As described above, we revised this sentence as "During two especially clean air events (in October and February) at Mount Zugspitze, the $\boldsymbol{\delta}^{13}C\text{-}CO_2$ was approximately -7‰, during which the CO_2 mixing ratios varied between 390 and 395 ppm."

L449. As with CO2, doing an afternoon hours-only analysis of model performance would be valuable.

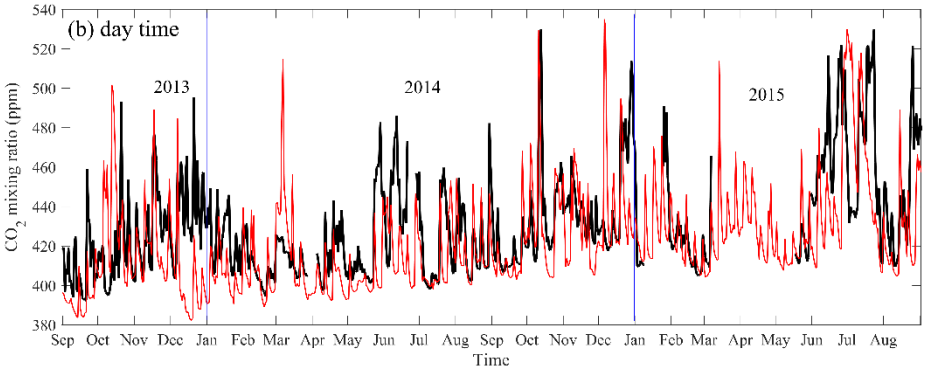

Figure R4b. Comparisons of CO_2 daily averages for day time.

As described above, we calculated the RMSE, R and MB for daily averages for all-day and daytime, which were 18.68 ppm, 0.44 and 2.68 ppm for all-day averages, and 25.21 ppm, 0.38 and 10.74 ppm for daytime.

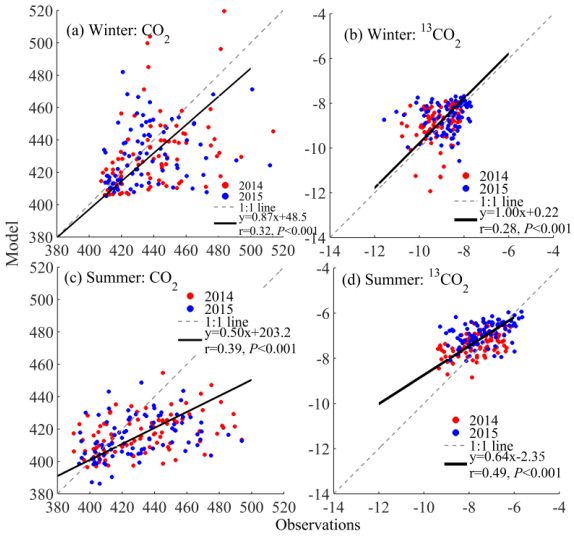

Figure R5. Scatter plots of observed versus modeled (a) winter time CO_2 mixing ratios, (b) winter time $\boldsymbol{\delta}^{13}C\text{-}CO_2$, (c) summer time CO_2, and (d) summer time $\boldsymbol{\delta}^{13}C\text{-}CO_2$ for both years, here these dots are day-time(10:-16:00) averages.

We also did the comparisons by only choosing daytime observations, the results indicated daytime CO_2 mixing ratio simulations in summer were slightly underestimated. While the simulations in winter can generally capture the trends for both CO_2, during which the biological CO_2 enhancement played a relatively small role than anthropogenic emissions. This larger bias in summer can potentially be caused by coarse spatial-temporal resolutions in Carbon Tracker biological CO_2 flux, which were 1×1 degree with three-hours averages, and urban biological CO_2 flux with finer spatial scale and hourly resolution is suggested in following studies. We added this description on lines 533-542.

L468. 'and generally caused by…' you could test this hypothesis by correlating the model minus obs. residuals for CO2 and δ13C. That is, if δ13C simulation errors are caused by CO2 errors, there should be a strong correlation and the slope should be related to the mean isotopic signature of the sources.

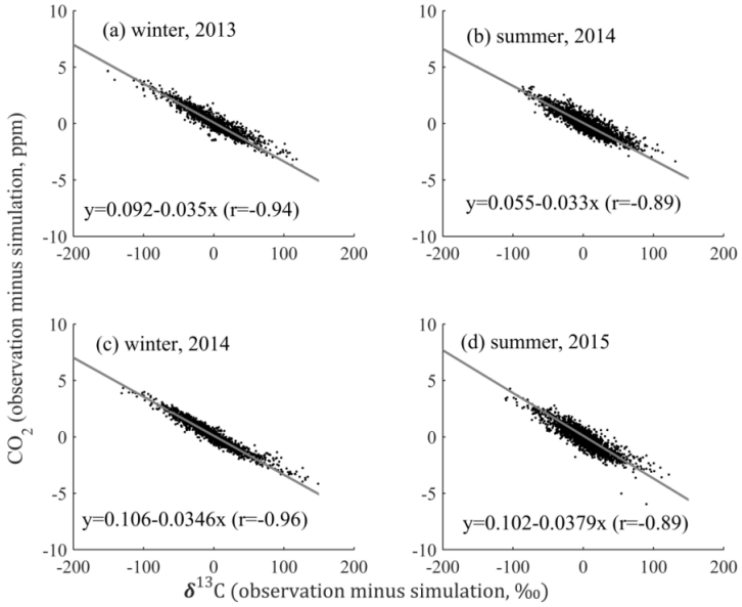

Figure R10. Relationship of observation minus simulation residual between CO_2 and $^{13}CO_2$ for (a) winter in 2013, (b) summer in 2014, (c) winter in 2014, and (d) summer in 2015.

Thanks for this suggestion. We have shown the relationship between observation minus simulation residual for CO_2 and $^{13}CO_2$ as displayed in the figure above. Strong negative correlation (with r<0.85) was observed. We have added this figure to the supplemental file and added text "Some large discrepancies are evident and generally caused by the simulated total CO_2 enhancement biases (potentially caused by poorly simulated PBLH during these periods) and the negative relationship between δ^{13}C-CO_2 and the CO_2 enhancement as shown in Figure S6." on lines 525-528.

L471. In Figure 9, by focusing only on daytime data you may minimize the impact of large night time NEE enhancements (Fig. 7b) and get a better correlation.

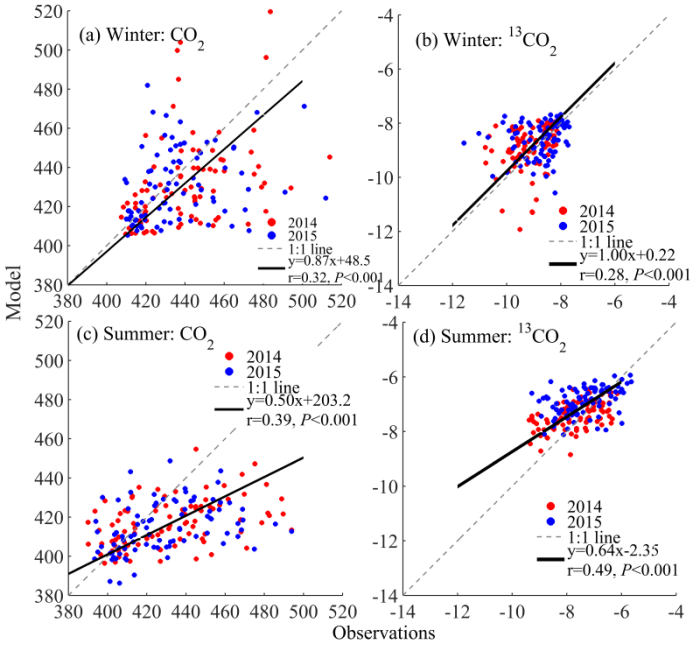

Figure R6. Scatter plots of observed versus modeled (a) winter time CO_2 mixing ratios, (b) winter time $\boldsymbol{\delta}^{13}C\text{-}CO_2$, (c) summer time CO_2, and (d) summer time $\boldsymbol{\delta}^{13}C\text{-}CO_2$ for both years, here these dots are day-time(10:00-16:00) averages.

Done as suggested, we have replied this in above questions.

L477. After 'was observed in December and July' reference the relevant figure.
We added "Figure 10a" following "December and July".

L484. Regarding delta_s, presumably this was calculated using equation 6 and the emission proportions listed on line 397. And are these the all-day values shown in Figure 12? Or are these data derived delta_s values from Keeling Plots. As mentioned earlier, different notation for simulated and data-based delta_s would be helpful.
The delta_s was simulated $\boldsymbol{\delta}_s$ (as described in equation 6) by using the simulated enhancement proportions. They are both the nighttime and all-day values in Figure b-c. We have revised as $\boldsymbol{\delta}_{s\_sim}$ and added "We define $\boldsymbol{\delta}_{s\_sim}$ as the simulated carbon isotope ratio of all sources to differentiate it from the observed $\boldsymbol{\delta}_{s\_obs}$" on line 303 for clarification.

L496. The seasonal cycle attribution is very confusing to me. To start with, Fig. 10 shows that about two thirds of the seasonal cycle is from the δ13C background. This seems straightforward. However, then you say that the background and regional source terms are 59% and 41% of the seasonal cycle. And then you say that total CO2 enhancement and CO2 enhancement components further contribute another 20% each. This seems to add up to 140%.

The regional source term contains both total CO_2 enhancement and CO_2 enhancement components, so here the contribution of 20% for each is the partition in regional source term, not the added contribution, we rephrased this sentence as "Further, the total CO_2 enhancement and change in $\boldsymbol{\delta}$s, sum of both can be treated as regional source term, contributed equally (about 20%) to the $\boldsymbol{\delta}^{13}C\text{-}CO_2$ seasonality." on lines 565-566 for clarification.

L499. I don't think your approach to separately investigate photosynthesis and respiration is valid. Negative NEE instances will still have a substantial influence from respiration. And some positive NEE will have substantial photosynthesis. I highly recommend simplifying the analysis (and make it more accurate) and only analyze the effect of NEE. You just don't have enough information to estimate the effects separately (unless perhaps you use the GPP estimates to partition NEE).
Thank you for these comments. We have revised this as follows: "(1) excluding negative NEE when photosynthesis is stronger than respiration, and (2) excluding both photosynthetic discrimination and respiration. Note that only NEE was used in our study with no partitioning between photosynthesis and respiration in the daytime. The only role of photosynthetic discrimination should be stronger than in case 1 when only negative NEE is used.". We also changed the caption and label from "without/excluding photosynthesis" to "without/excluding negative NEE" in Figure 10 for clarification.

L506-507 'via respiration' I would change this to 'via net respiration', or 'positive NEE', because photosynthesis might be active at some points.
Done as suggested, we changed 'via respiration' to 'via net respiration'.

L525. Eliminate 'a' prior to 0.40 per mil.
Done as suggested.

L526. Why is the impact of cement expressed as a range?
Here we used both observed $\boldsymbol{\delta}^{13}$C-CO_2 seasonality (1.51‰ and simulated $\boldsymbol{\delta}^{13}$C-CO_2 seasonality (1.53‰) to subtract the seasonality without non-metallic mineral production sources of 1.47‰. Here, we only use simulated $\boldsymbol{\delta}^{13}$C-CO_2 seasonality and revised "0.05‰ to 0.07‰" as "0.07‰"

L542. Another notation question: what exactly is delta_ms? It has not been defined.
We changed $\boldsymbol{\delta}_{s\_ms}$ with $\boldsymbol{\delta}_{s\_sim}$ and define it on line 303 as "We define $\boldsymbol{\delta}_{s\_sim}$ as the simulated carbon isotope ratio of all sources to differentiate it from the observed $\boldsymbol{\delta}_{s\_obs}$".

L561. It's not clear that earlier in the paper you stated a hypothesis.
We revise this sentence as "These results also indicated that".

L565. I disagree that the cement isotopic signature is the most distinct. In fact it is only ~ 8 per mil away from the atmospheric value, while biological sources and other fossil sources are ~ 20 per mil away and thus should exert more leverage on the atmospheric value. I also disagree that its emission is large. It is large for cement compared to other parts of the world, but it is still only ~ 10%.
We revised this sentence as "As discussed above, cement CO_2 emissions had the most distinct $\boldsymbol{\delta}^{13}$C-CO_2 end-member value of 0‰ ± 0.30‰ when compared with the averages of other anthropogenic sources. Combined with its large emission compared to other regions of the world, it had a strong potential to influence $\boldsymbol{\delta}_s$ and $\boldsymbol{\delta}^{13}$C-CO_2" on lines 634-637.

L567. Provide a reference for the statement that the YRD is the largest cement producing region in the world.
We added three references to support it as (USGS, 2014; Cai et al., 2015; Yang et al., 2017) on line 638.

L588. Regarding the 16.85 %, why is there such a large seasonal cycle in the cement enhancement proportion. Can you discuss this or offer any explanation? Seasonal changes in wind direction? Also, I am a bit confused about the a and b superscripts in Table S2. I thought the cement enhancements were simply calculated by convolving the cement emissions and the footprints. Do a and b just refer to whether the proportion is relative to total flux or only anthropogenic flux? If so, I would not use 'considering' in Table S2 and instead say explicitly that a is ratio of cement/anthro and b is a ratio of cement/total.
The large seasonal cycle of cement enhancement proportion should be driven by source area

changes as discussed "We found a relatively large difference between the enhancement proportion and the emission proportion for oil refineries (from 11.5% to 4.1%) as compared to other categories. This may be because power industry, manufacturing and non-metallic mineral production were more homogeneously distributed compared to oil refineries, which were closer to our CO_2 observation site. Further, changes in source footprint caused by wind direction variations likely played an important role." on line 441-446.
We revised the caption in Table S2 as "note the superscript 'a' is ratio of cement to anthropogenic CO_2 emissions and 'b' is a ratio of cement to total CO_2 emissions, which contains biological and anthropogenic CO_2 flux"

L609. I don't think you can conclude that delta_s is more sensitive to cement emissions than other emission categories without testing those categories!
We revised it as "These results indicate that $\boldsymbol{\delta}_s$ is sensitive to cement CO_2 emissions".

L616. 'contributed 0.40 per mil' to the seasonal cycle? Please say this explicitly.
We revised it as "contributed 0.40‰ to the seasonal cycle, accounting for 64.5% in all regional source terms (0.62‰)"

L631. Change 'write' to 'wrote'
Done as suggested.

L665. You shouldn't really cite a 'Discussions' paper from 2016. Please replace with the final paper (or remove if not available).
Done as suggested, we cited its formally published version.

L684. The journal title seems to appear twice.
Thanks for pointing it out, we deleted the first title.

L845. What is [J]?
We deleted [J].

L901. It would help to link panel B to panel A so one can see which of the cement locations are in which domain. Maybe draw some of the domain boxes in panel B?
Most of the sources came from domain 3, so we draw domain 3 in b.

*L931. Are the units really nmol/m2/s? This implies a *maximum* flux of 1e-14 mol/m2/s, based on the colorbar.*
The unite is nmol/m2/s, we have revised this typo.

L932. The colorbars/legends are very hard to see in the population density maps. Also, panel B colorbar needs units. Also for panel B, which domain in Figure 1 does this concentration footprint correspond to. It's not clear.
We increase the size of colorbar, we also added the units for panel B in the caption. The footprint corresponds to the same STILT domain setup in (Hu et al., 2019), we added it on

line 268.

L959. Consider using something besides 'Delta_CO2' as the y-axis label, because this has already been defined differently elsewhere in the paper. Also, in the figure caption be more clear about what 'Delta' means, which (I think) are differences between the same monthly averages in two different years. Panels b,c, and d are referred to after Fig. 5a and are not strongly related to panel a. I recommend separating them from this figure. Also what does 'distance xx' in the legend mean?

We changed added 'Delta_CO2' as "CO_2 enhancement difference," and revised the caption as "(a) Relation between monthly PBL height and change in CO_2 mixing ratio, here these dots represent difference of monthly averages in two different years for all hours;" we moved Figures 4b-d after Figure 5, and revised it caption as Time series (2013 to 2015) of (b) NDVI, (c) SIF, and (d) GPP. The distance indicates the radius of area centered with NUIST observation site, and the NDVI, SIF, GPP values are averages in these areas." in the caption of Figure 5.

L971. In panel a should 'model' really be 'anthropogenic'? b and c are switched in the caption.

Here the model represented the sum of both anthropogenic and biological CO_2 enhancement, we added "note 'model' represents the sum of both anthropogenic and biological CO_2 enhancement simulations," on line 1052. We also switched the caption for b and c.

L1019. Do the individual data points in the plots represent daily means? Please clarify.

Yes, these dots represent daily means. We added "here these dots are daily averages" on line 1097 for clarification.

L1051. Relabel panel a y-axis as delta_s.

Done as suggested.

Note. There are some language erros/typos in the supplement that need to be fixed

Done as suggested.