Review of Battle et al. "Atmospheric measurements of the terrestrial O_2:CO_2 exchange ratio of a mid-latitude forest."

Britton Stephens, National Center for Atmospheric Research, Boulder, CO, USA

Overview:

This is a fantastic and much needed data set and the authors have done a careful job on the analyses. What I find particularly compelling are the consistent and systematic differences in O_2:CO_2 ratios detected at night versus day, and within vs above the canopy, which likely will be of great use to ecosystem modelers tuning a new generation of ecosystem models that track O_2 and which are needed for investigating controls on α_B. This study is deserving of publication in ACP, and overall the presentation is good. However, I disagree with the broad extrapolation of the results to α_B and have several other substantive comments that I hope the authors are able to address.

Major comments:

1) Extrapolation to α_B

The parameter α_B in the global carbon budgeting exercise corresponds to the ratio of the global net O_2 and CO_2 exchanges associated with the *unknown* net land carbon sink and any other net land oxygen sources. Because it is the ratio of unknown processes, it is unknowable itself, and all we can do is refine our best guesses of it based on measurements of various pools and fluxes, such as those as presented here. The misrepresentation of measurements of ratios of specific pools or fluxes as measurements constraining the global net α_B already exists in the literature, so the authors are not the first to make this generalization, but I do not think it is justified. Instead, they should state explicitly what is measured and clearly indicate the assumptions that are involved in extrapolation to a global ratio of the net land sink.

More specifically, the measured ratios presented here represent either the local respiratory flux (night), the combined influence of local respiration and photosynthesis (day), or the combined influence of these processes and external (e.g. fossil fuel) influences. They do not represent the O_2:CO_2 ratio of the net carbon sink and oxygen source in Harvard Forest, nor of course globally. Indeed, on local scales if photosynthesis happens at a slightly different ratio than respiration and these two components are nearly in balance, any value (from – to + infinity) is possible for the ratio of the net exchanges. On global scales, we know the net carbon sink is far from zero, so its ratio to the net O_2 source is constrained to be close to that of the corresponding pool(s). However, if the net sink is a result of more leaves, more fine roots, more soil organic matter, or more stem wood, we would expect very different ratios locally, and values for α_B globally.

Especially because this paper identifies clear differences between day and night ratios, possibly indicating systematic differences between photosynthesis and respiration, it is important to be very careful when using either of these measurements, or their average, to estimate a ratio of net fluxes, locally or globally.

Fixing this oversimplification is largely a matter of wordsmithing, and I make specific suggestions here. However, the authors may prefer to use additional terms (e.g. photosynthetic quotient, respiratory quotient, or oxidative ratio) or to define new ones to explicitly describe what has been measured.

1.1) Specific suggestions to more accurately discuss relationships to α_B

Page 1 Lines 4-5, change “terrestrial photosynthesis and respiration” to “the net land sink”

Page 1 Lines 10-11, delete “, suggesting that this slope is our best estimate of α_B”

Page 1 Line 12, delete the last clause or change to something along the lines of “a more complete picture of the ratios for the component fluxes and potential pools for the net sink is needed.” Also, it would be good to call here or in the conclusion for ecosystem modeling that tracks O_2 and can explore potential values of α_B while matching observations.

Page 2 Line 2, change “terrestrial organic matter” to “the global net land carbon sink”

Page 2 Line 7, replace “effectively α_B” with “one contribution to α_B”

Page 2 Line 13, change “an average” to “the net imbalance”

Page 2 Line 17, add “of α_B” after “estimates”

Page 2 Line 20, change “An alternative approach” to something like “Another method that can inform on the processes responsible for α_B”

Page 2 Lines 23, change “a whole-ecosystem average value of α_B” to “the ratio for a particular flux component or mixture of components at a particular time for the ecosystem.”

Page 2 Line 25, add “contributions to” before “α_B”

Page 2 Line 31, change “land-ocean partition” to “ecosystem and process specific ratios”

Page 7 Line 17, change section heading to something like “the relationship between local O_2:CO_2 ratios and α_B”

Page 9 Line 9, change “α_{FF}” to “external fossil-fuel (numbers from 1.17 to 2.0) influences”

Page 9 Lines 12-14, change “in agreement with the values” to “similar to the stock-based estimate” and add “but is not a direct measure of the ratio of the net carbon sink at Harvard Forest or globally” to the end of the paragraph.

Page 10 Lines 4-6, we know the forest is not in balance, especially during summer but also on annual means, so this should not be considered just a possibility. Indeed, trying to estimate the O_2:CO_2 ratio of the small net imbalance is what makes these small differences important.

Page 10 Line 17, change “values for α_B” to “stock-based ratios”

Page 12 Lines 3-5, the sentence starting with “If our measurements. . . “ should be deleted, or rewritten to state something along the lines of “In the absence of fossil-fuel influence, our measured ratios correspond to local signatures of photosynthesis and respiration. The closeness of these ratios to 1 is consistent with measurements in other temperate forests but we do not yet have enough measurements to know if they are globally representative. Also, it is still possible that gross photosynthetic and respiratory exchanges could happen at ratios different from that corresponding to the much smaller annual net uptake of carbon and release of oxygen." You could also add "Nonetheless, our measurements provide valuable insight into the nature of O_2 and CO_2 exchange in a temperate forest and suggest that a better guess of the ratio to use in global carbon flux partitioning calculations, α_B, may be closer to 1 than previously used. Studies that rely on α_B should certainly explore sensitivity to values as low as 1.0 (e.g. Resplandy et al., Science, 2018)."

Page 12 Line 17, change "are in fact globally applicable" to "match the ratio of the global net carbon sink"

Then in all these places (and maybe a few I missed): Page 7 Line 22, Page 7 Line 24, Page 9 Line 9, Page 9 Lines 12, Page 9 Line 28, Page 10 Line 13, Page 10 Line 32, Page 10 Line 2, Page 12 Line 6, Page 12 Line 10, Page 12 Line 11, Page 12 Line 12, and Page 12 Line 15, change "α_B" to "local biotic [or photosynthetic or respiratory] flux ratios" or something similar.

2) Variability of ratios over years and seasons

These are discussed in Section 5.3 but not presented in either timeseries or seasonal cycle form. These temporal variations would help to evaluate the potential influence of fossil fuel emissions (which should have greater influence in winter), environmental drivers (is one year different from others and if so why?), and the repeatability of the ratios from day to day and year to year. I did not find the periodograms helpful in this regard and suggest replacing them with a timeseries figure of daily, weekly, or monthly ratios and also a composite seasonal cycle figure of either weekly or monthly mean ratios. It would also be helpful for at least 1 of the methods in Table 2, to show values calculated for summer and winter separately.

3) Influence of fossil fuel emissions vs respiration on day-night and low-high differences

I think that given the location of Harvard Forest, the influence of plumes of fossil-fuel pollution on the mean ratios deserves more discussion. As the authors point out in Section 5.3, the earlier study by Potasnak et al. (1999) shows fossil fuel to be a strong driver of variability in winter. Presumably this would have an effect on the all-season mean ratios presented here, but without a seasonal breakdown (see (2) above) this is difficult to assess. More specifically, I am wondering why the low ratios are systematically closer to 1 than the high ratios for both day and night and all calculation methods (Table 2). That the night-time ratios are systematically further than 1 than daytime can be explained by either respiration of high-N leaves and roots (e.g. Severinghaus, 1995 Fig. 4.1) or by greater fossil fuel influence. I think the latter is what the authors are referring to when they mention local versus regional influences in Section 5.2, and they point to fossil fuel influencing seasonal slope differences in Section 5.3, but the statements on Page 9 Lines 4-5, Page 11 Line 30 seem to discount it.

The day-night and low-high differences are really interesting, so clarifying their influences would be helpful. Specifically, one might think that the night time and lower level data would be more strongly influenced by respiration, and that respiration might have O_2:CO_2 ratios further from 1, as speculated by Severinghous (1995). However, while the observations do show ratios further from 1 at night as might fit this pattern, they show ratios closer to 1 from the low intake, implying more fossil influences at the high level. The authors sort of state this but it could be stated more explicitly that the ratios for the night-low intake are the best estimates of a purely respiratory signature. Their differences with the day-low ratios is suggestive that the picture of Severinghaus (1995, Figure 4.1) may apply, but their range (-0.99 to -1.08) depending on method is suggestive of less N oxidation in the in situ / modern Harvard Forest soil than Severinghaus (1995) found in his older samples analyzed in lab.

Additional text (or a barplot version of Table 2) might help to highlight the observed differences. Also, it may be possible to further isolate local biotic signatures by calculating ratios of differences between the high and the low inlet at night, as remote fossil fuel influences likely affect both heights similarly (the greater fossil influence at the high intake is because of reduced biotic influence not because of differences in fossil influence). Was this tried and if so what did it show?

Minor comments and suggestions:

Page 1 Line 10: clarify that this number is the average and standard error of 6-hour periods.

Page 1 Line 10: Why do you expect biotic influences will dominate during the day? With photosynthesis and respiration in opposition, daytime CO_2 fluctuations are generally smaller than at night, when respiration is unopposed and mixing is less vigorous.

Page 1 Line 31: can you say any more in the paper about “the response of the ecosystem to environmental controls” based on interannual variations in ratios or seasonal variations (see (2) above).

Page 2 Line 10, insert “and whether N is oxidized to nitrate” after “respired”

Page 3 Line 17, if the filters are outside the traps as shown (and presumably warm) how do they trap ice crystals?

Page 3 Line 30, does it really take 4 valves to make the changeover? It would be helpful to see the individual valves in the schematic to see how this works.

Page 5 Line 12, change “lines” to “regulators” unless you have reason to specify otherwise – permeation through elastomers in the regulator is typically the reason for needing to purge. It also would be helpful to specify the regulators used.

Page 5 Line 15, it would be helpful to know when (seconds after switch) typically 70% of the change has happened. I also wonder how much the remaining 30% of rollover is affecting the signal. Did you try 80% or 90% here?

Page 5 Line 20 and Page 6 Line 26, I don’t think “S2” is the official name of the scale. Please check with Ralph Keeling but I think something like the “Scripps O_2 Program O_2 Scale” is more appropriate. Also, specify WMO X2007 CO_2 scale (if indeed that is the scale used). Finally, missing space after “scale”

Page 5 Lines 30-31, How does a leak affect precision – variable fractionation or variable room air contamination? After tightening the packing nut did the problem go away? A leak like that could have temperature sensitive fractionation, potentially affecting ratios – did the ratios during this period differ from before?

Page 6 Line 2 and elsewhere, I agree with Reviewer 1 that O_2 values should not be reported in umol/mol. Rather they should be reported in per meg, and ratios calculated after converting O_2 to "ppm equivalents".

Page 6 Line 6, the dominant source of variation about the fit line is likely not due to instrumental imprecision but rather real atmospheric variability from multiple sources that violate the simple model of a linear fit. Thus, a better scaling for the Deming regression would be something like 1:1 on a molar basis (or up to 1.4 to 1 if fossil fuel was the source of poor fit). This should at least be tested to see if it affects the results.

Page 6 Line 16, is the scatter from run to run not just instrumental drift over 6 hours?

Page 6 Line 31, insert "stronger influence of" before "soil"

Page 7 Lines 18-21, please clarify what is meant by "this conceptual framework generally holds." Since 1.4 is only the average fossil fuel ratio and coal at 1.17 and natural gas at 2.0 vary considerably, I would expect more than 2 end members most of the time. How variable is the O_2:CO_2 ratio for different pollution events at Harvard Forest? Also, I suggest citing Keeling dissertation (http://bluemoon.ucsd.edu/publications/ralph/34_PhDthesis.pdf) and COFFEE (https://doi.org/10.5194/acp-11-6855-201) dataset for further information of fossil fuel ratios.

Page 8 Line 3, please explain the rationale for using ½ the PBL height. Assuming fluxes mixed over the full PBL height would seem more appropriate. I don't think I've managed to fully grasp this from the Lin reference and the personal communication cited is not helpful.

Page 8 Line 14, I think this is as expected from the set up of the calculation – a better test would be to see what happens when the occasional plume of natural gas burning arrives at the site (or coal, or ocean influence). This would probably require using LPDM footprints and a spatially explicit fossil fuel emission map, or better yet a 3D transport simulation. Since both of these are likely beyond the scope here, please just acknowledge the limits to the conclusions that can be drawn here.

Section 4.2, it's not clear to me where the results of this analysis are shown. A map of average surface influence would be more helpful than a single example. Also, is the height of the particles relative to the PBL height used to calculate the region of influence? I do not think 6 particles per receptor is enough, and something like an LPDM with 100s to 1000s of particles per receptor would be more appropriate to the task. Also, since the trajectories here are only run for 6 hours I don't think their spatial range can be used to define the region of influence – locations several days back will also have influence. Without footprints defined by vertical particle locations, where to cut this off is not well defined and I don't think the statement regarding consistency with Gerbig et al. is justified.

Page 8 Line 32, what specific region is their estimate for and how did you convert this to the region around Harvard Forest – are they the same and it's just a unit conversion, or was some other downscaling needed?

Page 9 Lines 4-7, I wonder if comparing to fossil fuel influences on the basis of regional average emissions captures all of the potential impact given that fossil fuel emissions would often arrive at the site in concentrated plumes. It might not take very many plumes of natural gas emissions at -2.0 O_2:CO_2 to affect the average ratio, if they are strong. This should be mentioned or discussed.

Page 9 Lines 11-12, some discussion and justification of the 3-sigma cutoff seems warranted. What are sigma and SEM before the cut (and perhaps add this info to Table 2)? What types of events is this cutoff meant to exclude (instrument problems, pollution plumes, low CO_2 variability, or all of the above) and does their exclusion sway the average ratios? You might also consider filtering based on low CO_2 variability to exclude ratios calculated when the denominator is small, which may be more agnostic on the actual ratios allowed past the filter.

Page 9 Line 15, I think “Further confirmation . . . considering” should be replaced with “We also considered” as I don’t think the appropriateness of this model has been confirmed and there are other ways to get ratios between 1 and 1.4 with many end members.

Page 9 Line 21 and elsewhere, somewhere please define what you mean by “steeper” and “shallower” slopes as these terms can be ambiguous.

Page 9 Lines 29-32, I think the text from “The problem with” to the end of the paragraph needs to be deleted or revised to account for the example illustrated by Severinghaus (1995), especially his Figure 4.1 in which different ratios for photosynthesis and respiration can be a permanent feature given the flux of N from leaves to soil via litter (see (3) above).

Page 10 Line 2 and elsewhere, it would also be helpful if you clarified that “larger” ratios really means more negative.

Page 10 Lines 4-8. I find the assertion that respired carbon is on average as young as weeks (here and in the Wehr and Saleska 2015 reference) somewhat puzzling. We know that about half of respiration is from soil organic matter that had to grow and die so this number must be much greater. The cited reference appears to use a general correspondence of seasonal cycles in the ^{13}C ratio of photosynthesis and respiration to make this claim, but this seems a bit tenuous and ^{14}C provides a much better estimate of carbon age. Trumbore et al. (Ecol. App. 2000, https://doi.org/10.1890/1051-0761(2000)010[0399:AOSOMA]2.0.CO;2) use ^{14}C to show that Harvard Forest soil respiration (root respiration + decomposition) has an average age of 3 years. Even allowing for a 5-25% contribution from stem respiration, this is considerably longer than the weeks mentioned here. Please try to reconcile these 2 approaches and edit the text accordingly.

Page 10 Lines 14-16, I agree with these statements but think they are indicating a detectable fossil-fuel influence on the high intake ratios and the high-low differences. It would be good to clarify this here and also acknowledge it when discussing fossil-fuel influences earlier (see (3) above).

Page 10 Line 20, please clarify how “stability might explain the shallow day-low slopes”? Is this because of less fossil-fuel influence, more photosynthesis influence, or something else? One might expect more respiration influence low and that this might lead to steeper, not shallower slopes so it would be good to spell the argument out.

Typographical comments and suggestions:

Many places in manuscript, including the title, there appears to be an extra space after subscripts.

I believe there should be spaces between all values and their units.

I believe units should not be italic.

Page 1 Line 9: use same number of significant digits for both subsets

Page 2, Line 1, insert “global” before “average”

Page 2 Line 6, insert “global” before “carbon”

Page 2 Lines 18-19 use “$PgCyr^{-1}$”

Page 3 Line 4, say “5 m” here but “6 m” in abstract, please clarify

Page 5 Line 7, add “seconds” after “120” (if that is what is meant), also probably don’t need quotes around live since already used without.

Page 7 Line 27, add “we” before “estimate” and change “fluxes” to “contributions”

Page 8 Line 2, 22.4 corresponds to 0 C – there may be a better value or values to use here

Page 10 Line 23, “a” not “an”

Page 11 Line 12, “influences” after “anthropogenic”

Page 12 Line 3, change “molar ratio” to “all-data average molar ratio over 6-hr periods”

Page 13 Line 6, add “the” before “period”

Table 2: present the 4 metrics in the same order in each section

Figure 3: It is not possible to distinguish the 2 colors used