Peer Review File

Manuscript Title:

Amazonia as a carbon source linked to deforestation and climate change

Editorial Notes:

Reviewer Comments & Author Rebuttals

Reviewer Reports on the Initial Version:

Ref #1

Comments on "Decrease in Amazonia carbon uptake linked to trends in deforestation and climate" by L. Gatti et al

This is one of the most ambitious and potentially impactful manuscripts I've ever reviewed. The authors report results of nearly 600 vertical profiles of trace gas measurements made from small aircraft across the Amazon forest over 9 years. It's an unprecedented data set and shows distinct and provocative patterns in space and time. The authors highlight the difference between small carbon losses in the eastern Amazon and small carbon gains in the western Amazon, especially due to fires in the dry season. Using climate and land use data, they propose a mechanism to explain these variations as a result of seasonal warming and drying as well as deforestation. The results will impact our understanding of the carbon balance of tropical forests, the vulnerability of their carbon stocks, the dependence of terrestrial carbon sinks on climate change, and the future trajectory of carbon-climate feedback in the Earth system.

Unfortunately, the manuscript is diffuse and vague with respect to key methods. The authors underestimate critical uncertainties, and fail to distinguish clearly among hard data, inference, and speculation. This study certainly can and should be published but the paper must be substantially revised to focus clearly on what was done, what was found, and what it means so that readers can evaluate for themselves whether the conclusions are supported by the observations.

The data presented represent an incredible, even heroic effort sustained across a decade. To my knowledge there are no comparable data over tropical forests anywhere else, and this is the largest tropical forest on Earth. Complementary literature reports on forest inventories, eddy covariance towers, ecosystem manipulations, remote sensing, and numerical simulation. These other methods point to a huge and vulnerable stock of stored terrestrial carbon in the region which may or may contribute to strong positive feedback on global warming. Forest inventory sampling, eddy covariance, and ecosystem manipulations are severely limited by their very small spatial scales. Remote sensing can't yet directly assess small changes in stocks. Numerical modeling is less grounded in data and results are all over the map. Climate scientists desperately need results like those reported here, which report direct measurements of atmospheric variations across a huge domain in space and time. It's worth a substantial effort to work with the authors to get this paper published.

Problems with the Narrative

The results show a clear difference between carbon cycling in the eastern vs central and western Amazon, especially during the long and increasingly hot/dry dry season, with stronger atmospheric CO_2 drawdown over the west than the east. The authors interpret these results as being caused by fires as well as by increasing physiological stress due to rising temperatures in the east relative to the west. Furthermore, they suggest that deforestation and fire in the east is responsible for much of the warming and drying in that region.

All of this may well be true, and if so the chain of causality that may link land use to fires to regional amplification of climate change and a change from carbon sink to carbon source in tropical forests is one of the most important stories of global change in the 21st Century. But it is absolutely crucial that the authors distinguish between reporting their unique data (especially the long-term CO_2 and CO concentration profiles) and their somewhat less solid interpretation of the complex of linked mechanisms that controls variations of these atmospheric gases. The manuscript as it stands mixes experimental results, discussion, conclusions, and in some places speculation in such a way as to distract from the key contribution that this study can provide to our understanding of ecosystem-climate feedback in the Earth system.

Part of the narrative problem with the present manuscript may be a consequence of the unique format of Nature articles in which the body of the paper is expected to read almost like an extended abstract or a "News and Views" article about the study, and most of the supporting material is buried in extended material which many readers may not even examine in detail. In this case, there's an additional layer of methodological detail which must be mined from a series of recent publications in Atmosphere and other journals. Few readers will reach that deeply to understand exactly what has been done in the current study and will instead just read the main article. This criticism is not aimed at the authors but rather the journal, but it is imperative that the paper be structured to highlight its most important results.

A revised manuscript should first present the data collected in the field, note the differences in column trace gas gradients between eastern and western profiles, and pose testable hypotheses about the mechanisms explaining them. The rest of the manuscript should be framed as a series of tests of the hypothetical mechanisms: simple atmospheric budgets, correlations with climate trends and land use, etc. A concluding paragraph is the place for clearly-identified speculation about the implications of the findings of this study for our expectations of the response of Amazon carbon cycling to potential future warming and drying.

The paper should include a concise summary of the broader context of the study with quantitative Amazon carbon budget estimates from previous results of forest inventory, remote sensing, inverse modeling, and ecosystem modeling (using consistent units). It is not necessary to be exhaustive here; the point is that previous studies wildly disagree. That's one of the key motivations for the manuscript and highlights the need for the complementary approach afforded by the aircraft sampling program described here.

Problems with Atmospheric Budgets

Unlike other regions, the tropical forest is characterized by strong uptake and release of CO_2 in *all* seasons, so net CO_2 exchange is much weaker and the regional/seasonal differences are more subtle than in midlatitudes. This makes it more difficult to quantify surface carbon budgets from atmospheric measurements than other regions. The authors address this signal:noise problem by aggregating huge numbers of measurements to obtain robust results for climatological gradients over quarters and years. Furthermore, they have paired climatological gradients of CO_2 and CO to try to separate the relative contributions of fire and ecosystem fluxes on the carbon balance of each region. These methodological choices are commendable, and this study represents by far the best characterization of CO_2 gradients ever obtained in the tropical atmosphere.

The primary results of this study are the time-mean differences (horizontal gradient) in vertically-integrated CO_2 and CO concentration at each profiling site relative to a reference concentration derived from background sites thousands of kilometers upwind in the South Atlantic Ocean. The primary result of the paper is that these differences show that CO_2 (and CO) are added to the atmosphere during the dry season as air flows across the Amazon from the coast, especially at the eastern sites. By contrast the atmospheric column is depleted of CO_2 during the wet season and is approximately neutral at the western sites. These results are incredibly solid. The measurements themselves are made in a state-of-the-science lab by some of the world-class experts in the experimental techniques. The calibration and quality control is excellent. Everything else in the paper (trajectory modeling, atmospheric budgets, remote sensing, climate data analysis) is in some sense secondary analysis intended to interpret these primary results.

The first and arguably most important secondary analysis in the study is to convert atmospheric concentration gradients to quantitative carbon budgets using a very simple mass balance approach. The ocean background concentration is subtracted from the vertically-integrated mass of tracer (CO_2 or CO) at the sampling site and then divided by the transit time from ocean to profile based on a simple online trajectory model (equation M1, line 398). Several questionable assumptions underly this calculation:

* The lower troposphere over the region flows coherently as a “slab” for a period of several days from “background” stations to profile sites and can be characterized by a single “transit time”;

* The three island/coastal surface sampling stations (or a linear mixture of them) represent the same 4.4 km deep atmospheric column sampled by the aircraft profiles; and

* There is no horizontal or vertical exchange between the layer of air from the surface to 4.4 km altitude and the deeper layer above it (which itself may or may not have had contact with the surface) during the transit from background stations to profile samples.

The HYSPLIT model used in the study differs from derivates such as HYPACT in that it diagnoses only back trajectories and does not attempt to diagnose the influence of the surface by counting the number of “surface impacts” of air parcels due to turbulence or convective updrafts/downdrafts. Assuming that the HYSPLIT trajectories are in fact perfectly accurate, they can’t say whether air flowed across a region in the atmospheric boundary layer (in contact with the surface) or aloft. In fact over the course of a day or two the air certainly takes many paths to reach a given location.

The authors have done some reasonable calculations to assess random error in their flux estimates: subsampling the data, time shifting the trajectories, propagation of sampling variance and uncertainties in the column means. The atmospheric transport works in favor of their simple mass balance in that steady deep Trade Winds from the coast toward the west are the norm. The meteorology is in this sense far simpler than a temperate location disturbed by fronts and synoptic events.

The tropical troposphere is frequently disturbed by strong vertical motions due to convective updrafts and downdrafts, sometimes reaching three times the height to which the small aircraft used to collect air samples can fly. Convective events are never sampled by the aircraft (for safety reasons), but cumulus convection between the coast and the sampled columns introduces systematic error into the simple atmospheric budget represented by equation M1. This type of systematic error can’t be assessed by the statistical methods described under “Monte Carlo error propagation” in lines 573-584.

If we could treat unknown convective mixing as a simple loss term that dilutes the influence of surface fluxes on the measured trace gas column, the authors might apply a multiplicative factor $C > 1$ in M1 to account for convective loss from the column en route.

The trouble is that we have no way to properly constrain the value of C given the data collected by the sampling program. Surely the uncertainty in the value of this loss term is the leading term in the uncertainty in the calculated surface fluxes. With some detailed meteorological analysis, perhaps a case might be made that the uncertainty in C is less than a factor of two, but I can’t see how the very low flux uncertainties (order 0.1 or even 0.01 gC m^{-2} day^{-1}) reported in the main body of the manuscript can be justified.

For reference to other studies of carbon budgets in the Amazon a net flux with precision 0.01 gC m-2 day-1 is equal to a time mean eddy flux of less than 0.01 micromoles m-2 sec-1 or a forest inventory with precision better than 0.04 tons C ha-1 yr-1. No flux tower or inventory study could possibly achieve anything close to that level of precision, so the fluxes reported from aircraft profiling could never be calibrated against data on the ground.

If we could be confident that convective column losses of surface flux impacts on measured gas concentrations are fairly consistent, then the important conclusions of the paper with respect to regional and seasonal differences in fluxes may be robust. For example, the critical results summarized in Figure 4 might remain qualitatively identical if C=2 in every case, but with different numbers on each arrow. The authors' interpretation of these regional differences as resulting from a combination of deforestation, drought, and fires (all shown to be more prominent in the eastern rather than western Amazon) is consistent with the concentration data in any case, based on the spatial patterns.

The worst case for uncertain convective losses involves systematic seasonal dependence of C, which might impact source estimates more than sink estimates or fire fluxes more than biogenic fluxes, for example. If this were the case, then linear sums of biased fluxes would not be expected to add to annual carbon budgets, and our confidence in even the qualitative spatial patterns of dry season flux components would become suspect. In a revised manuscript, the authors must seriously consider whether seasonally or meteorologically dependent convective losses might have such an effect on their results and provide evidence that they don't.

Please understand that I am not suggesting that the simple atmospheric budgets presented in the manuscript are wrong or that there is some technical error in them which should be "fixed." Rather I would argue that the way these budgets are reported in the manuscript is unrealistically precise. More importantly, the precise budgets are not important to the main results of the study! The critical result is the difference in lower tropospheric concentration gradients in the east vs the west. The precise numerical values of mass balance budgets derived from these gradients provides a convenient metric for comparing them, but is not critical to the interpretation and the paper loses credibility by overstating the confidence with which these rough numbers are known.

The key contributions of the study are
1) The collection of many hundreds of vertical profiles of measured lower tropospheric CO2 and CO over nearly a decade spanning a huge area of the Amazon;

2) The documentation of significant differences in concentration gradients between the eastern and western Amazon, especially in the dry season; and

3) The association of those differences with land use, fire, and secular trends in climate

These three pillars can stand even if the metric of quantitatively precise flux estimates is compromised by the possibility of influence of surface fluxes on column concentration due to uncertain convective loss to the upper troposphere.

I strongly suggest that the authors consider submitting a revised manuscript which focuses on the three key points above, which can be made to some degree independently of the precisely quantitative flux estimates.

Detailed Comments on the Manuscript

Line 59: consider changing "farming" to "agriculture" and "agriculture" to "crops"

Line 60: consider changing "was" to "has been"

Fig 1: The definition and especially the visual indication of the region of influence is not clear from the caption. Is the ROI indicated by the light blue perimeter? How is this line defined?

Line 129: “Climate change is promoting a reduction in precipitation” It seems to me that the reduction in precipitation is climate change, by definition. One could argue that a change in global circulation or some other aspect of global change is responsible for the change in precipitation, but “climate change” is too amorphous for this statement.

Line 162: These fluxes are are all quite small and indicate that the forest is very close to carbon balance (much smaller than some of the results of global inverse or ecosystem models). Some context from previous literature would be helpful to the reader. Also as I’ve mentioned I think the reported uncertainties are much too small.

Lines 161-165: On one hand, this paragraph is reporting a primary result (flux components and totals) for the northeast region. But the very same text insinuates that deforestation and strong seasonality are to blame for carbon losses from the region via physiological stress. This is a complex attribution argument. It needs careful explication and should be separated from the basic summary of experimental results. This paragraph is a great example of the confusion of results, interpretation, conclusions, and speculation that I find most problematic in the manuscript.

Fig 3: If I’m right and the uncertainty bars here are underestimated, then the simple conclusion is that all four regions are close to carbon balance and that the imbalance and its variability is predominantly controlled by biomass burning.

Lines 177-178: The claim that ecosystems in the region are a net source of CO_2 (not counting fire) is very weakly supported by the data, especially if the flux uncertainties are underestimated as I’ve argued above

Lines 182-187: “NBE changed from a sink to a strong source …” Actual detection of flipping a huge region from a carbon sink to a source as a result of drought stress is a very strong claim and must be evaluated carefully in the context of previous studies. The plot certainly does look like a trend, but Nyears is small and the estimated fluxes are still very close to carbon balance. Do the linear correlations ($r=0.68$) include uncertainty in the estimated atmospheric budgets? Would they still be statistically significant if the uncertainties were twice as large?

Lines 205-209: You are associating strong seasonality with dry season stress and with annual carbon balance. This theme is an important part of your overall interpretation and might be called out in the abstract/intro/conclusions. It’s not clear a priori to a reader unfamiliar with the region that for example a strongly seasonal EVI signal would be associated with carbon losses or why. It’s worth making this argument explicitly and clearly.

Lines 230-232: The estimated carbon balances here are unrealistically precise and in any case pretty small.

Lines 235-237: This is the central assertion of the paper. Be very careful here to distinguish between “what the data show” and “what we think they imply”

Line 238-239: There is a further insinuation that “human disturbance” (deforestation & forest degradation) have CAUSED the changes in dry season climate, which are in turn CAUSING physiological stress for intact forests in the east and therefore causing the whole region to switch from a carbon sink to a source in the annual mean. This is NOT DIRECTLY SUPPORTED by the data. It’s a perfectly reasonable story but the manuscript MUST be revised to clearly differentiate among data, hypotheses, and interpretation/conclusions

Lines 244-256: This entire paragraph is interpretive speculation. It’s fascinating and provocative, but isn't a result. That's perfectly OK but please be careful to distinguish between results and interpretations

Line 246: consider adding the word "could" "deforestation and degradation in the east [could?] make the area more susceptible to fire"

Lines 257-263: Besides Hubau et al, this concluding paragraph would be a good place for some more context with other studies of whole Amazon carbon budgets. Are the confidence intervals here (and throughout this manuscript) 1-sigma or 2-sigma estimates?

Lines 386-389: Why average the concentration first and then subtract the mean background? Wouldn't the reverse order make more sense? Annual averages of the difference from the background?

Line 401 (eq M1): It appears that each air parcel is assumed to be (1) in contact with the surface and (2) not in contact with any air above 4.4 km asl during the entire transit time. No horizontal or vertical divergence? No updrafts or downdrafts except in the lower troposphere?

Line 414: The treatment of the three Atlantic sites (as far away as Cape Town!) is reasonable but it's counterintuitive that weekly samples at the surface from so far away can be treated as inflow air that can be quantitatively compared to an entire profile over the interior Amazon to derive ecosystem carbon balances over enormous regions. The details must be dug out of the more detailed papers (eg ref 39), so the average reader of the main article is likely to be a bit misled about this critical assumption. This is an example of why I think the uncertainties (1-sigma?) in the derived carbon budgets are underestimated.

Lines 460-465: This is another source of my unease in the extremely precision of estimated mass balance reported throughout the main manuscript. It appears from the text here and in ref 41 that ROI is the cumulative density of HYSPLIT trajectories *passing through* a grid cell in a calendar quarter. There appears to be no treatment of the altitude at which those trajectories cross? The only way this corresponds to a surface influence function is if the air is vertically mixed? This is not unreasonable, but the assumptions are aggressive and contribute structural uncertainty to the flux estimates which has not been addressed in the Monte Carlo methods described in the body of the manuscript or in lines 576-587

Ref #2

A: Summary of the key results

This paper uses a unique (not reported elsewhere) aircraft profiling dataset over Amazonia, covering nearly a decade from 2010. It shows that the eastern, and especially south eastern, Amazon has become a net C source to the atmosphere, driven by a combination of human disturbance and climate change.

B: Originality and significance

While not surprising, the quality of the dataset and narrowness of the error bars make the study results highly significant and novel. Being able to pick apart the fire emissions from the forest net biome exchange figures is a really useful factor of this method, and the long time series with consistent data enables trends to be seen.

The specific climate findings (lines 103-109), while not the focus of the data collected in this study, also struck me as fascinating. While I have seen some of these data before, I have not seen the results but in such a striking way, nor such a strong impact highlighted: they show the 3 dry season months have experienced warming of 1.37 degrees over Amazonia, with rainfall in these months down 17% - compared to 1 degree of warming and no change in precipitation for the whole year. This strong seasonal stress if vital to understanding the findings in the paper, and has wide consequences. It seems the Amazon dry season is experiencing worse climate change that the average, with similar devastating consequences as the polar regions are experiencing. And the highlighting of this being more extreme in the SE site – 2.5 degrees higher in the dry season, is shocking – unsurprising the forest Net Biome Exchange (excluding fire) is positive here, driven by climate change.

C: Data & methodology: validity of approach, quality of data, quality of presentation

The sample size (600 flights from 4 sites) is very large, and sufficient to draw the conclusions the authors offer. The Regions of Influence calculations appear robust, though this is not my specific area of expertise so I cannot comment on the appropriateness of the method use.

In general I find the methods logical and well explained, and the approach robust.

Some figures are a bit difficult to read, or don't show the data to its full potential impact – e.g. Fig. 2 where the text is small and the panels hard to compare.

D: Appropriate use of statistics and treatment of uncertainties

Uncertainties are well covered – the Monte Carlo approach followed appears well justified, and results are mostly presented with confidence intervals that are reasonably calculated.

Lines 182-184, using a simple multivariate linear correlation to explain the important change in NBE from sink to source is a bit unsatisfying, if not strictly wrong. The number of data points may not allow for anything more significant, but it would be good to explain which of these (temperature or water storage) was stronger, and if there was any interaction in the two. And an explanation should be included as for why GRACE data was used rather than VPD or precipitation – I'd like the authors to present the data to check they've not just chosen the variable with the highest correlation, but then not explained it's lower for the others.

We're left to take the author's word that TAB (2010-12) and TEF (2013-18) are similar enough to be analysed as one trend. This looks reasonable from the map shown in Figure 1, and the evidence of no discontinuity in Extended Data figures 6 and 7c – but at the very least a line Extended Data Fig.6 showing where the two sites switched over would be useful. And I think it would be useful to have a specific figure/table covering this issue, giving the average ROI covered by each, %historical deforestation of each, and how these change with the time of year – to prove this is acceptable.

It is a shame there is so much missing data, but the length of the time series and the 'MissForest' approach taken to gap fill seem to be okay. I don't think these gaps, if filled with real data, would change any of the conclusions.

E: Conclusions: robustness, validity, reliability

I have no doubt that the conclusions are robust, valid and reliable. The conclusions come directly from the data collected, and the implications are self-evident.

While alluded to a little at the end, I do have a slight worry that the NBE figures are treated as comparable to the fluxes due to intact forest, e.g. as compared to the Hubau et al. (2020) paper in the final paragraph of the main text. These fluxes do not just contain intact forest – they contain many areas of Human Modified Forest, including logging concessions actively undergoing logging and recovering from past logging, and forests undergoing recovery from other forms of disturbance (e.g. last year's fires) – plus non-forest vegetation will be contributing to these figures too. This is not a problem with the data or analysis – the facts and figures stand – but I would like more notes of caution from the authors around these areas, especially to encourage people to be careful of comparisons.

F: Suggested improvements:

Small point, but neither reference 7 or 8 give information on 'below-ground organic carbon'. 7 does include a belowground biomass layer, but adding that to the AGB gives a total value for Amazonia nearer 100 Pg than 200 Pg C. Their number sounds about right for a figure including soil carbon, but they need to fix the references.

Lines 126-127 – it is strange to give so many (useful and interesting) facts and figures about temperature rises in the dry season for each site – but not to give figures for precipitation and VPD changes, instead just referring to the literature. I think it would be helpful for the reader to have these listed, or summary figures could be added to Fig. 2 to assist with comparing between panels? I know these might duplicate those in Fig. 4, so maybe this isn't necessary – but Fig. 4 only has the ASO figures, not those for the whole year. VPD seems important too, was it not possible to quantify changes in this over the sites?

The units throughout the results are given in gC m-2 d-1. This is a perfectly acceptable unit given what is measured, but is hard to compare to the fluxes given in Pg carbon, e.g. in the first sentence of the main paper, and in the discussion of carbon sinks/sources here and elsewhere (e.g. line 240). I wonder if units of Pg C yr-1 could be easier to compare, though they would need to be standardised by area somehow so might get unwieldy.

Line 528-529, I would have liked to have seen information as to how the Burned Area product was resampled to 1x1 degree resolution. The original 500m data are binary (the whole pixel is burned on a particular calendar day or not). Was the ultimate 1 degree product, as I assume, the proportion of that 1 degree pixel that burned within a particular time period (e.g. 3-month period)? I assume this has been done properly, but it should be clearer in the text.

Lines 531-542, would be good to have justification for why EVI was used, rather than say NDVI, and why given this is about (according to this paragraph) fPAR, the explanation then talks about 'greater biomass' – fPAR and biomass are not well correlated (nor are EVI and biomass beyond a certain quite low level). I don't think anything has been done incorrectly here, but I think a bit more thought should be shown as to what EVI is showing (definitely something about leaves more than biomass – the seasonal trends seen in EVI in the paper are not seasonal biomass trends!) and why it is chosen (lower artefacts than NDVI typically through inclusion of blue band).

There are climate models that make predictions about carbon fluxes (including and without fire) at a sub-basin scale. The authors will certainly be aware of these, with several involved in intercomparison papers – it would be interesting for the reader to know if in general terms the eastern vs western Amazon are behaving as models would expect given this level of climate change an atmospheric CO2 concentration, or whether these are outwith the expectations of the majority of models. This may be too big a separate piece of analysis, but if relatively trivial to do and explain I think it would add a useful 'implication' paragraph in the main text of the paper, either validating the models we are using, or suggesting they are not performing well in a critical region.

G: References: appropriate credit to previous work?

Relevant papers appear to have been cited.

H: Clarity and context: lucidity of abstract/summary, appropriateness of abstract, introduction and conclusions

I believe the paper is well written overall, with some smaller problems as highlighted above.

Ref #3

The manuscript presents an evaluation of the Amazon carbon balance, using air column measurements at 4 sites with a basins of observation that covers substantial portions of the Eastern and Central Amazon. The key observations are 1) that the Amazon has been a net carbon source from 2000-2018 and 2) that fire and deforestation have made the Eastern Amazon more of a source than the western Amazon. The dataset is rich and seemingly definitive, and it is on the merits of the empirical work that the manuscript merits consideration for publication in Nature.

Personally, I see three sites acting as a sink, and only one site in the SouthEast acting as a source (and not the entire Eastern Amazon as currently framed). Overall, the data and trends seem quite logical and consistent with the published literature to me, although I would reframe the summary of the results.

On that subject, I'm surprised to see so little context w.r.t. what is known about Amazon-wide carbon fluxes especially about fire, since the authors argue that fire is driving source vs. sink outcomes. Fire in the Amazon is quite well studied but mostly neglected as a point of discussion in the manuscript. I can find only one publication (#10) that deals with fire fluxes in the entire paper, which is quite thin given its central importance as a driver of basin-wide carbon dynamics (in all regions, frankly).

Overall, I found the manuscript somewhat difficult to read. They rely heavily on site acronyms in the results section, which requires a reader to go back and memorize which is which, so this really impedes clear communication. The results and discussion should really be reorganized if the target is a short format journal – instead of presenting the drivers and carbon results separately, these should be integrated into a region-by-region combined results and interpretation. Then end with a substantial section evaluating the effects of different drivers on spatial and temporal trends in carbon fluxes across the basin.

Fig 1 but also a larger methodological point: it's convenient that the site areas match up exactly with the Amazon mask. Obviously this is not how it actually works – ALF for instance might include a significant non-forest component in the atmospheric signal, so too RBA (whereas San and TEF presumably include a lot of ocean in the signal which may also be an issue?). In the case of RBA, some substantial component of the fire signal may be coming from the cerrado, so that may be misattributed here. This needs to be dealt with directly, at least in the figure but more likely in the method overall.

Line 37: Even a few words about what the framework, instead of just calling it "simple"
Line 40: remove "relatively pristine" – unnecessary value judgement
Line 43-44: something funny about the sentence construction here. A flux can't be a source, although a region can be. Rephrase.

Line 53: "around 2200 mm per year" This is a weird number to put here, since it varies substantially across the basin. Either remove entirely, or add "on average, across the basin, rainfall around 2200 mm per year"

Line 112: It's quite an empty statement to say that 2 out of only 4 stand out. Are you defining the mean based on the other two?!

Line 155: On a semantic level, it's counterintuitive to me that fire is not a part of total flux (although obviously it isn't a part of NPP). This may be standard terminology, although I have not encountered it, but at a minimum I would be sure to include definitions of these in the captions for figures where fluxes are presented. It's a real schlep to sift through the text to find a definition in order to interpret a figure.

Author Rebuttals to Initial Comments:

Response to Reviewers

Dear Dr. Juliane Mossinger and Reviewers

Thank you for your useful comments and suggestions on the language, structure and clarifications of our manuscript.

We believe that the advice received was of great importance to improving the quality of our manuscript, such that it could be published in a renowned journal, such as Nature.

Since the changes at the article promoted a very extensive change in the structure, it was not possible to left all changes highlighted. But at each answer we provided a detailed information about the new localization (lines) at this revised version.

We apologize for the delay in sending the response considerations proposed, but we made new improvements with data analysis, recalculated the uncertainty, and promoted the requested changes in the text.

We have modified the manuscript accordingly, and detailed corrections are listed below point by point. We believe the paper is now more adequate for publication.

We look forward to your response.

Thank you again for the all comments.

Best Regards,

Luciana V. Gatti and the co-authors.

Referees' comments:

Referee #1 (Remarks to the Author):

Comments on “Decrease in Amazonia carbon uptake linked to trends in deforestation and climate” by L. Gatti et al

1- This is one of the most ambitious and potentially impactful manuscripts I’ve ever reviewed. The authors report results of nearly 600 vertical profiles of trace gas measurements made from small aircraft across the Amazon forest over 9 years. It’s an unprecedented data set and shows distinct and provocative patterns in space and time.

The authors highlight the difference between small carbon losses in the eastern Amazon and small carbon gains in the western Amazon, especially due to fires in the dry season. Using climate and land use data, they propose a mechanism to explain these variations as a result of seasonal warming and drying as well as deforestation. The results will impact our understanding of the carbon balance of tropical forests, the vulnerability of their carbon stocks, the dependence of terrestrial carbon sinks on climate change, and the future trajectory of carbon-climate feedback in the Earth system.

Unfortunately, the manuscript is diffuse and vague with respect to key methods. The authors underestimate critical uncertainties, and fail to distinguish clearly among hard data, inference, and speculation. This study certainly can and should be published but the paper must be substantially revised to focus clearly on what was done, what was found, and what it means so that readers can evaluate for themselves whether the conclusions are supported by the observations.

The data presented represent an incredible, even heroic effort sustained across a decade. To my knowledge there are no comparable data over tropical forests anywhere else, and this is the largest tropical forest on Earth. Complementary literature reports on forest inventories, eddy covariance towers, ecosystem manipulations, remote sensing, and numerical simulation. These other methods point to a huge and vulnerable stock of stored terrestrial carbon in the region which may or may contribute to strong positive feedback on global warming. Forest inventory sampling, eddy covariance, and ecosystem manipulations are severely limited by their very small spatial scales. Remote sensing can’t yet directly assess small changes in stocks. Numerical modeling is less grounded in data and results are all over the map. Climate scientists desperately need results like those reported here, which report direct measurements of atmospheric variations across a huge domain in space and time. It's worth a substantial effort to work with the authors to get this paper published.

Authors answer: We would like to thank you for your so careful, profound and very profitable analysis that will make our article much better and clearer. The paper structure was modified to make clear separation between the finds from gradient results and fluxes, climatological analysis, contrast east vs west and the interpretation for the results and finally our hypothesis.

2- Problems with the Narrative

The results show a clear difference between carbon cycling in the eastern vs central and western Amazon, especially during the long and increasingly hot/dry dry season, with stronger atmospheric CO2 drawdown over the west than the east. The authors interpret these results as being caused by fires as well as by increasing physiological stress due to rising temperatures in the east relative to the west. Furthermore, they suggest that deforestation and fire in the east is responsible for much of the warming and drying in that region.
All of this may well be true, and if so the chain of causality that may link land use to fires to regional amplification of climate change and a change from carbon sink to carbon source in tropical forests is one of the most important stories of global change in the 21st Century. But it is absolutely crucial that the authors distinguish between reporting their unique data (especially the long-term CO2 and CO concentration profiles) and their somewhat less solid interpretation of the complex of linked mechanisms that controls variations of these atmospheric gases. The manuscript as it stands mixes experimental results, discussion, conclusions, and in some places speculation in such a way as to distract from the key contribution that this study can provide to our understanding of ecosystem-climate feedback in the Earth system.

Part of the narrative problem with the present manuscript may be a consequence of the unique format of Nature articles in which the body of the paper is expected to read almost like an extended abstract or a "News and Views" article about the study, and most of the supporting material is buried in extended material which many readers may not even examine in detail.

Authors answer: We produced changes in the manuscript to make clearer the presentation of results, we add a section named "Atmospheric carbon vertical profiles" from line 68 to 120 (640 words) and added new figures: Figure 2, Extended data Figure 1 and Supplementary Information Figure 1. We calculated the gradients using 2 methods: vertical profiles less background (BKG) and the top of vertical profiles. A study was made comparing the top of vertical profile and BKG to observe the losses of information we don't capture in our method by convections systems in the atmosphere (Supplementary Information Figure 1). After this section is presented the flux results, in sequence the climatological analysis, the comparison east vs west and finally the interpretation and hypotheses.

To help in use the extended data and supplementary information we will choose open access paper for publication

3- In this case, there's an additional layer of methodological detail which must be mined from a series of recent publications in Atmosphere and other journals.

Authors answer: We added a methodological resume using Supplementary information bringing the most important lack of information.

4- Few readers will reach that deeply to understand exactly what has been done in the current study and will instead just read the main article. This criticism is not aimed at the authors but rather the journal, but it is imperative that the paper be structured to highlight its most important results.

A revised manuscript should first present the data collected in the field, differences in column trace gas gradients between eastern and western profiles, and pose testable hypotheses about the mechanisms explaining them. The rest of the manuscript should be framed as a series of tests of the hypothetical mechanisms: simple atmospheric budgets, correlations with climate trends and land use, etc.

Authors answer: We appreciate the suggestions that bring more strong information based on observations. The new version of paper explores more the vertical profiles CO_2 concentrations. It was added a figure to the main paper (Figure 2) showing the annual mean vertical profiles and the mean for all years and the figure Extended data Figure 1a shows mean below boundary layer (lower than 1.5 km) and top of vertical profiles (higher than 3.8 km) for all sites and all vertical profiles and the table with the results at Extended data Figure 1d. The Figure 2 shows clear evidence the CO_2 contribution from surface to the atmosphere, based on data observation (we add also a table with the mean concentration: Extended data Figure 1c).

5- A concluding paragraph is the place for clearly-identified speculation about the implications of the findings of this study for our expectations of the response of Amazon carbon cycling to potential future warming and drying.

The paper should include a concise summary of the broader context of the study with quantitative Amazon carbon budget estimates from previous results of forest inventory, remote sensing, inverse modeling, and ecosystem modeling (using consistent units). It is not necessary to be exhaustive here; the point is that previous studies wildly disagree. That's one of the key motivations for the manuscript and highlights the need for the complementary approach afforded by the aircraft sampling program described here.

Authors answer: We appreciate the suggestions in include the comparison with others approaches. The discussion about the Amazon Carbon Balance and previous studies for comparison now is presented from line 173 to 188. It is clear the discrepancy from bottom-up approach since forest plots at mature forests and the top-down studies like satellites and vertical profiles, that capture all the process: sinks and sources.

6- Problems with Atmospheric Budgets

Unlike other regions, the tropical forest is characterized by strong uptake and release of CO2 in *all* seasons, so net CO2 exchange is much weaker and the regional/seasonal differences are more subtle than in midlatitudes. This makes it more difficult to quantify surface carbon budgets from atmospheric measurements than other regions.

The authors address this signal:noise problem by aggregating huge numbers of measurements to obtain robust results for climatological gradients over quarters and years.

Furthermore, they have paired climatological gradients of CO2 and CO to try to separate the relative contributions of fire and ecosystem fluxes on the carbon balance of each region. These methodological choices are commendable, and this study represents by far the best characterization of CO2 gradients ever obtained in the tropical atmosphere. The primary results of this study are the time-mean differences (horizontal gradient) in vertically-integrated CO2 and CO concentration at each profiling site relative to a reference concentration derived from background sites thousands of kilometers upwind in the South Atlantic Ocean.

Authors answer: We improved the analysis based in CO_2 concentrations, comparing the mean concentrations below boundary layer (<1.5 km height) with the top of vertical profiles (> 3.8 km height). This analysis avoids any problem with BKG methods. The BKG methods used here was a result of 2 years study and 7 methods tested. The background is related of the position / concentration in the virtual line RBP-ASC and starting in this line, at longitude 30 degree, going down straight to South Pole. It was observed a very small gradients below equator, between ASC and CPT (at Extended Data Fig 1a ASC: medium grey, CPT: dark grey). The main air circulation trajectories in this part of South Atlantic Ocean go in the direction to South Pole. This is the reason for the virtual line at 30 longitude. We add a supplementary Information Figure 5 detailing this method.

7- The primary result of the paper is that these differences show that CO2 (and CO) are added to the atmosphere during the dry season as air flows across the Amazon from the coast, especially at the eastern sites. By contrast the atmospheric column is depleted of CO2 during the wet season and is approximately neutral at the western sites. These results are incredibly solid. The measurements themselves are made in a state-of-the-science lab by some of the world-class experts in the experimental techniques. The calibration and quality control is excellent. Everything else in the paper (trajectory modeling, atmospheric budgets, remote sensing, climate data analysis) is in some sense secondary analysis intended to interpret these primary results.

The first and arguably most important secondary analysis in the study is to convert atmospheric concentration gradients to quantitative carbon budgets using a very simple mass balance approach. The ocean background concentration is subtracted from the vertically-integrated mass of tracer (CO2 or CO) at the sampling site and then divided by the transit time from ocean to profile based on a simple online trajectory model (equation M1, line 398). Several questionable assumptions underly this calculation:

* The lower troposphere over the region flows coherently as a "slab" for a period of several days from "background" stations to profile sites and can be characterized by a single "transit time";

* The three island/coastal surface sampling stations (or a linear mixture of them) represent the same 4.4 km deep atmospheric column sampled by the aircraft profiles; and
* There is no horizontal or vertical exchange between the layer of air from the surface to 4.4 km altitude and the deeper layer above it (which itself may or may not have had contact with the surface) during the transit from background stations to profile samples.

<u>Authors answer</u>: To address this uncertainty in the flux calculation we used the difference between the top of vertical profile (>3.8 km) and BKG as an uncertainty associated with BKG, because these differences may represent the losses along the air masses pathway. For example, for RBA this difference was 1.9 ppm (Supplementary Information Table 1).

8- The HYSPLIT model used in the study differs from derivates such as HYPACT in that it diagnoses only back trajectories and does not attempt to diagnose the influence of the surface by counting the number of "surface impacts" of air parcels due to turbulence or convective updrafts/downdrafts. Assuming that the HYSPLIT trajectories are in fact perfectly accurate, they can't say whether air flowed across a region in the atmospheric boundary layer (in contact with the surface) or aloft. In fact over the course of a day or two the air certainly takes many paths to reach a given location.

The authors have done some reasonable calculations to assess random error in their flux estimates: subsampling the data, time shifting the trajectories, propagation of sampling variance and uncertainties in the column means. The atmospheric transport works in favor of their simple mass balance in that steady deep Trade Winds from the coast toward the west are the norm. The meteorology is in this sense far simpler than a temperate location disturbed by fronts and synoptic events.

The tropical troposphere is frequently disturbed by strong vertical motions due to convective updrafts and downdrafts, sometimes reaching three times the height to which the small aircraft used to collect air samples can fly. Convective events are never sampled by the aircraft (for safety reasons), but cumulus convection between the coast and the sampled columns introduces systematic error into the simple atmospheric budget represented by equation M1. This type of systematic error can't be assessed by the statistical methods described under "Monte Carlo error propagation" in lines 573-584.

Authors answer: To address a more realistic time trajectory uncertainty we used 3 different trajectory models. We compare HYSPLIT back-trajectories times for the year 2010 with two different models, FLEXPART Lagrangian particle dispersion model and back trajectories derived from the meso-scale model BRAMS. This comparison showed larger uncertainties than our previous estimation. Then, we consider the largest difference in mean profile travel time from HYSPLIT and the other two models using the RMSE values as back-trajectories time uncertainties (Supplementary Information Table 1).

9- If we could treat unknown convective mixing as a simple loss term that dilutes the influence of surface fluxes on the measured trace gas column, the authors might apply a multiplicative factor $C > 1$ in M1 to account for convective loss from the column en route. The trouble is that we have no way to properly constrain the value of C given the data collected by the sampling program. Surely the uncertainty in the value of this loss term is the leading term in the uncertainty in the calculated surface fluxes. With some detailed meteorological analysis, perhaps a case might be made that the uncertainty in C is less than a factor of two, but I can't see how the very low flux uncertainties (order 0.1 or even 0.01 gC m-2 day-1) reported in the main body of the manuscript can be justified. For reference to other studies of carbon budgets in the Amazon a net flux with precision 0.01 gC m-2 day-1 is equal to a time mean eddy flux of less than 0.01 micromoles m-2 sec-1 or a forest inventory with precision better than 0.04 tons C ha^{-1} y^{-1}. No flux tower or inventory study could possibly achieve anything close to that level of precision, so the fluxes reported from aircraft profiling could never be calibrated against data on the ground.

If we could be confident that convective column losses of surface flux impacts on measured gas concentrations are fairly consistent, then the important conclusions of the paper with respect to regional and seasonal differences in fluxes may be robust. For example, the critical results summarized in Figure 4 might remain qualitatively identical if C=2 in every case, but with different numbers on each arrow. The authors' interpretation of these regional differences as resulting from a combination of deforestation, drought, and fires (all shown to be more prominent in the eastern rather than western Amazon) is consistent with the concentration data in any case, based on the spatial patterns.

The worst case for uncertain convective losses involves systematic seasonal dependence of C, which might impact source estimates more than sink estimates or fire fluxes more than biogenic fluxes, for example. If this were the case, then linear sums of biased fluxes would not be expected to add to annual carbon budgets, and our confidence in even the qualitative spatial patterns of dry season flux components would become suspect. In a revised manuscript, the authors must seriously consider whether seasonally or meteorologically dependent convective losses might have such an effect on their results and provide evidence that they don't.

Please understand that I am not suggesting that the simple atmospheric budgets presented in the manuscript are wrong or that there is some technical error in them which should be "fixed." Rather I would argue that the way these budgets are reported in the manuscript is unrealistically precise. More importantly, the precise budgets are not important to the main results of the study! The critical result is the difference in lower tropospheric concentration gradients in the east vs the west. The precise numerical values of mass balance budgets derived from these gradients provides a convenient metric for comparing them, but is not critical to the interpretation and the paper loses credibility by overstating the confidence with which these rough numbers are known.

Authors answer: We appreciate the suggestions and many additional studies were done. The comparison between the top of vertical profile (> 3.8km) and BKG, shows a positive loss during the dry season, when fires formed Pyrocumulus clouds promoting strong mixing process in the atmosphere. On the other hand, we can also observe a negative loss during wet season, mainly at SAN (March, April and May) and RBA (January, February and March), see these results at Supplementary Information Figure 1.

To address a more realistic uncertainty to the calculated fluxes we revised all uncertainty sources (Supplementary Information Table 1):

a) For BKG we used the difference from top of vertical profiles (> 3.8 km height) and BKG, that addressed uncertainty from BKG calculation itself and the losses during the pathway with convective process in the atmosphere.
b) For the Time trajectory uncertainty, we consider the largest difference in mean profile travel time from HYSPLIT and the other two models (Flexpart and BRAMS) using the RMSE values.
c) We also add the CO flux uncertainty to NBE and Fire flux uncertainty, additionally to all others uncertainties sources. For total CO flux (F_{CO}) we estimate them as the CO_2 total flux and for the ($F_{CONatural)}$ we consider 25% of uncertainty
d) We also change the 9 years mean uncertainty: the theoretical uncertainty for the nine-year mean fluxes is [sigma = sqrt(sum(sigma_i^2))/9], but this assumes that annual fluxes are uncorrelated. To be conservative, allowing for significant year to year correlation, we calculate the nine-year uncertainties as [sigma = sum(sigma_i)/9]

All these changes promote an increase at all uncertainties, reflecting at the 9 years mean C total flux uncertainty for each site:

ALF: first version: 0.02 gC m^{-2} d^{-1} now: 0.09 gC m^{-2} d^{-1}

SAN: first version: 0.06 gC m^{-2} d^{-1} now: 0.25 gC m^{-2} d^{-1}

RBA: first version: 0.02 gC m^{-2} d^{-1} now: 0.11 gC m^{-2} d^{-1}

TAB_TEF: first version: 0.02 gC m^{-2} d^{-1} now: 0.08 gC m^{-2} d^{-1}

10- The key contributions of the study are the collection of many hundreds of vertical profiles of measured lower tropospheric CO2 and CO over nearly a decade spanning a huge area of the Amazon; The documentation of significant differences in concentration gradients between the eastern and western Amazon, especially in the dry season; and the association of those differences with land use, fire, and secular trends in climate. These three pillars can stand even if the metric of quantitatively precise flux estimates is compromised by the possibility of influence of surface fluxes on column concentration due to uncertain convective loss to the upper troposphere.

I strongly suggest that the authors consider submitting a revised manuscript which focuses on the three key points above, which can be made to some degree independently of the precisely quantitative flux estimates.

Authors answer: We appreciated the suggestions and produced many changes in the manuscript that improve its quality and structure. We believe now the paper is clear and the uncertainties are more realistic. Many thanks for the valuable suggestions and the critical review.

Detailed Comments on the Manuscript

11- Line 59: consider changing “farming” to “agriculture” and “agriculture” to “crops”

Authors answer: It was done. This phrase now is at line 60.

12- Line 60: consider changing “was” to “has been”

Authors answer: It was done. This phrase now is at line 61.

13- Fig 1: The definition and especially the visual indication of the region of influence is not clear from the caption. Is the ROI indicated by the light blue perimeter? How is this line defined?

Authors answer: It was added in the Figure 1 legend and was add the explanation in Methods, in the item Regions of Influence (The topic Regions of Influence is from line 600 until 624). The border definition is described in the lines: 615-618.

14- Line 129: “Climate change is promoting a reduction in precipitation” It seems to me that the reduction in precipitation is climate change, by definition. One could argue that a change in global circulation or some other aspect of global change is responsible for the change in precipitation, but “climate change” is too amorphous for this statement.

Authors answer: Thank you. The phase was really not clear and was modified for a better understand and now this part is from lines 226 until 228.

15- Line 162: These fluxes are all quite small and indicate that the forest is very close to carbon balance (much smaller than some of the results of global inverse or ecosystem models). Some context from previous literature would be helpful to the reader. Also, as I've mentioned I think the reported uncertainties are much too small.

Authors answer: The uncertainties changes was described above. The results are reported in gC m^{-2} d^{-1}. But If we use gCO_2 m^{-2} d^{-1} this value increase in a factor of 3.667.

For example, If the whole Amazonia (area adopted in this study 7.25 x 10^6 km^2) presented the same flux of SAN, the emission from Amazonia will be 4.0 Pg CO_2 y^{-1}. This amount represents 67% of mean total global land-use change CO_2 emission, according Global Carbon Budget 2020. In the other hand, if the whole Amazonia presents the same flux of RBA, the emission will be 0.4 Pg CO_2 y^{-1}. This amount represents 6,7% of mean total global land-use change CO_2 emission. This comparison shows significant differences among the studied sites.

We add also comparisons with others studies in the lines: 178 – 189.

16- Lines 161-165: On one hand, this paragraph is reporting a primary result (flux components and totals) for the northeast region. But the very same text insinuates that deforestation and strong seasonality are to blame for carbon losses from the region via physiological stress. This is a complex attribution argument. It needs careful explication and should be separated from the basic summary of experimental results. This paragraph is a great example of the confusion of results, interpretation, conclusions, and speculation that I find most problematic in the manuscript.

Authors answer: We restructured completely the manuscript.

Now:

a) Vertical profile data and studies based in CO_2 concentrations.
b) Flux results with Column budget technique and in the last part the C balance
c) Climate analysis for the last 40 years (temperature and precipitation) and deforestation for each region of influence.
d) Comparison east versus west in Amazon regions
e) Hypotheses about the link between the carbon fluxes and balance results and the human impact at it regions with deforestation and changes at dry season.

17- Fig 3: If I'm right and the uncertainty bars here are underestimated, then the simple conclusion is that all four regions are close to carbon balance and that the imbalance and its variability is predominantly controlled by biomass burning.

Authors answer: We change the Figure 3 with new uncertainties bars.

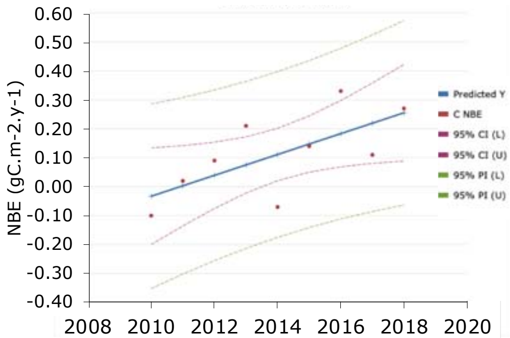
18- Lines 177-178: The claim that ecosystems in the region are a net source of CO2 (not counting fire) is very weakly supported by the data, especially if the flux uncertainties are underestimated as I've argued above

19- Lines 182-187: "NBE changed from a sink to a strong source …" Actual detection of flipping a huge region from a carbon sink to a source as a result of drought stress is a very strong claim and must be evaluated carefully in the context of previous studies. The plot certainly does look like a trend, but Nyears is small and the estimated fluxes are still very close to carbon balance. Do the linear correlations (r=0.68) include uncertainty in the estimated atmospheric budgets? Would they still be statistically significant if the uncertainties were twice as large?

Authors answer 18 and 19: Since this topic is about the same issue with the previous question, we will answer together.

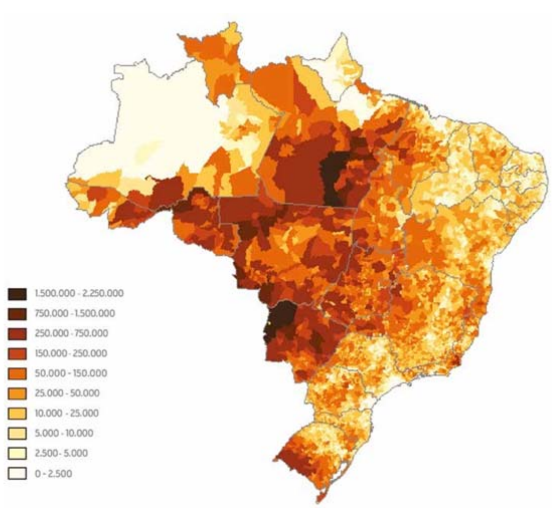
The phrase "NBE changed from a sink to a strong source" was changed in the lines: 146 to 149. This conclusion is supported by correlations between NBE and year, calculated by 3 different methods: Pearson method's = 0.688; Kendall method's = 0.555 and Spearman method's = 0.717.

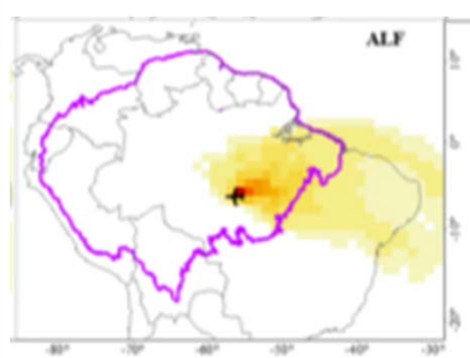
It is important to highlight each year mean is composed by 12 months that is composed by 1 or 2 vertical profiles and each vertical profile comes from 12 flasks. Since the mean by year vertical profiles (VP) is around 20 VP, each year mean is composed by 240 samples. Then 9 years is composed by 2160 samples, which can be considered a robust population. We worked with the annual mean to avoid the seasonality that could shows a false correlation. Then the conclusion that each year the NBE is higher is supported by statistical analysis. Below is the Figure with the confidence's interval of NBE vs year. Beside the figure the complete statistical analysis using the Statplus software.



Regression Statistics										
R		0.67654	R-Squared		0.45771	Adjusted R-Squared		0.38024		
MSE		0.01316	S		0.11472	MAPE		76.15102		
Durbin-Watson (DW)		2.71389	Log likelihood		7.84761					
Akaike inf. criterion (AIC)		-1.29947	AICc		-1.23598					
Schwarz criterion (BIC)		-1.25564	Hannan-Quinn criterion (HQC)		-1.39405					
PRESS		0.13578	PRESS RMSE		0.12283	Predicted R-Squared		0.20079		
C NBE = - 72.39289 + 0.03600 * year										
ANOVA										
	d.f.	SS	MS	F	p-value					
Regression	1	0.07776	0.07776	5.90824	0.04537					
Residual	7	0.09213	0.01316							
Total	8	0.16989								
	Coefficients	Std Err	LCL	UCL	t Stat	p-value	H0 (5%)	VIF	TOL	Beta
Intercept	-72.39289	29.82862	-142.92636	-1.85942	-2.42696	0.04562	Rejected			
year	0.036	0.01481	0.00098	0.07102	2.43069	0.04537	Rejected	1	1	0.67654
T (5%)	2.36462									
LCL - Lower limit of the 95% confidence interval										
UCL - Upper limit of the 95% confidence interval										

The uncertainty bar was included in the Extended Data Figure 5a and still show the consistent increase. This region is under strong human impact influence. A report from “Associação Brasileira das Indústrias Exportadoras de Carnes Bovinas (ABIEC) (Brazilian Association of Beef Exporters Industry)”. The Pará state show the city with the highest herd of cattle (Sao Felix do Xingu) at ALF region of influence. Another 11 cities at Para state from center to south show also are included in the report among the 40 cities with highest number of Cattle from Brazil. The map from 2018 IBGE (Brazilian Institute of Geography and Statistics) at left side and the ALF region of influence in the right side.





20- Lines 205-209: You are associating strong seasonality with dry season stress and with annual carbon balance. This theme is an important part of your overall interpretation and might be called out in the abstract/intro/conclusions. It’s not clear a priori to a reader unfamiliar with the region that for example a strongly seasonal EVI signal would be associated with carbon losses or why. It’s worth making this argument explicitly and clearly.

Authors answer: Thanks for your comment. The summary and main text were modified to address clearly this association (lines 45 to 48 and 247 to 271)

The text was modified and included some description for EVI at lines 135-137.

The presence of a strong seasonality provided by the EVI reflects an evident dry season, which is the former driver of increased fire emissions, as stated in the paragraph before. The EVI is an indirect measure of vegetation greenness (vigor), with the highest value ~1 representing the highest fraction of Photosynthetically active radiation (fPar) between 400-700 nm absorbed by plants, and low values ~0, the absence of fPar absorbed by plants. The low EVI in a dry season reflects a period of susceptible drier biomass for ignition, although it means nothing about water stress directly. Water stress is more related to an abnormal year or reduced water availability. We inserted a paragraph to explained it in a few words: *“EVI seasonality is also the smallest of all sites, which expresses a relatively high constant fraction of photosynthetically active radiation (fPAR) absorbed by plants, i.e., greater vegetation greenness and a lower fraction of dry biomass throughout the year, reducing the fire risk.*”

21- Lines 230-232: The estimated carbon balances here are unrealistically precise and in any case pretty small.

Authors answer: The changes at uncertainty was presented in previous answers.

22- Lines 235-237: This is the central assertion of the paper. Be very careful here to distinguish between “what the data show” and “what we think they imply”

Authors answer: The structure of the manuscript was changed and the new structure was presented in the answers of the questions 9 and 16.

23- Line 238-239: There is a further insinuation that “human disturbance” (deforestation & forest degradation) have CAUSED the changes in dry season climate, which are in turn CAUSING physiological stress for intact forests in the east and therefore causing the whole region to switch from a carbon sink to a source in the annual mean. This is NOT DIRECTLY SUPPORTED by the data. It’s a perfectly reasonable story but the manuscript MUST be revised to clearly differentiate among data, hypotheses, and interpretation/conclusions

Authors answer: We appreciated the suggestions and produced many changes in the manuscript that improve its quality and structure. We believe now the paper is clear and the uncertainties are more realistic. Many thanks for the valuable suggestions and the critical review. Now it is clearer our hypothesis.

24- Lines 244-256: This entire paragraph is interpretive speculation. It's fascinating and provocative, but isn't a result. That's perfectly OK but please be careful to distinguish between results and interpretations

Authors answer: Many thanks for the valuable suggestions. Now our hypothesis it is more clear.

25- Line 246: consider adding the word "could" "deforestation and degradation in the east [could?] make the area more susceptible to fire"

Authors answer: The phrase was modified and is presented at lines 277 to 284

26- Lines 257-263: Besides Hubau et al, this concluding paragraph would be a good place for some more context with other studies of whole Amazon carbon budgets. Are the confidence intervals here (and throughout this manuscript) 1-sigma or 2-sigma estimates?

Authors answer: The comparison was changed and is showed at lines 184 to 188. The comparison was done with a mean of 3 important papers from Rainfor project. It was made clear a statement that our NBE is not comparable with NEE, because NBE is composed by forest uptake, and all carbon sources, natural and anthropogenic (non-fire), like decomposition, that is consequence of deforestation and degradation promoted by fire and selective logging. The new uncertainties are 1- sigma, but compared with the old uncertainty at least 4-sigma.

27- Lines 386-389: Why average the concentration first and then subtract the mean background? Wouldn't the reverse order make more sense? Annual averages of the difference from the background?

Authors answer: Thank you in point out this issue. The phrase was not clear. The new explanation is at lines 529 to 531.

28- Line 401 (eq M1): It appears that each air parcel is assumed to be (1) in contact with the surface and (2) not in contact with any air above 4.4 km asl during the entire transit time. No horizontal or vertical divergence? No updrafts or downdrafts except in the lower troposphere?

Authors answer: The integration of vertical profile is a simple approach and to address a more realistic uncertainty to the calculated fluxes we change the uncertainty to address this issue in the BKG uncertainty the difference from top of vertical profiles (>3.8 km height) and BKG, which include the uncertainty from BKG calculation itself and the losses during the pathway with convective process in the atmosphere.

29- Line 414: The treatment of the three Atlantic sites (as far away as Cape Town!) is reasonable but it's counterintuitive that weekly samples at the surface from so far away can be treated as inflow air that can be quantitatively compared to an entire profile over the interior Amazon to derive ecosystem carbon balances over enormous regions. The details must be dug out of the more detailed papers (eg ref 39), so the average reader of the main article is likely to be a bit misled about this critical assumption. This is an example of why I think the uncertainties (1-sigma?) in the derived carbon budgets are underestimated.

Authors answer: The BKG methods used here was a result of 2 years study and 7 methods tested. The background is related of the position / concentration in the virtual line RBP-ASC and starting in this line, at longitude 30 degree, going down straight to South Pole. It was observed a very small gradients below equator, between ASC and CPT (at Extended Data Fig 1a ASC: medium grey, CPT: dark grey). The main air circulation trajectories in this part of South Atlantic Ocean go in the direction to South Pole. This is the reason for the virtual line at 30 longitude. We add a supplementary Information Figure 5 detailing this method, where can be found a comparison between our BKG and vertical profile performed at coast site (SAH). The correlation between BKG calculated by our method and CO_2 concentrations measured at vertical profiles coast site SAH was r^2: 0.77. At the Domingues et al (2020) paper was also compared the BKG calculated by SF6 method (Figure 2 from Domingues at al) with this new approach used here. It is also compared the BKG calculated by our method with the CO2 concentration at the 3 global stations (Figure 3 from Domingues at al).

30- Lines 460-465: This is another source of my unease in the extremely precision of estimated mass balance reported throughout the main manuscript. It appears from the text here and in ref 41 that ROI is the cumulative density of HYSPLIT trajectories *passing through* a grid cell in a calendar quarter. There appears to be no treatment of the altitude at which those trajectories cross? The only way this corresponds to a surface influence function

is if the air is vertically mixed? This is not unreasonable, but the assumptions are aggressive and contribute structural uncertainty to the flux estimates which has not been addressed in the Monte Carlo methods described in the body of the manuscript or in lines 576-587
<u>Authors answer</u>: That is true. The region of influence does not consider vertical mixing, such as "footprint" from FLEXPART, for instance. However, in the ref 41, we have tested the region of influence comparing two years footprint with our method and have proved that the area of surface cover by both is similar, regardless of considering altitude and method (i.e., FLEXPART or HYSPLIT). See the figure below

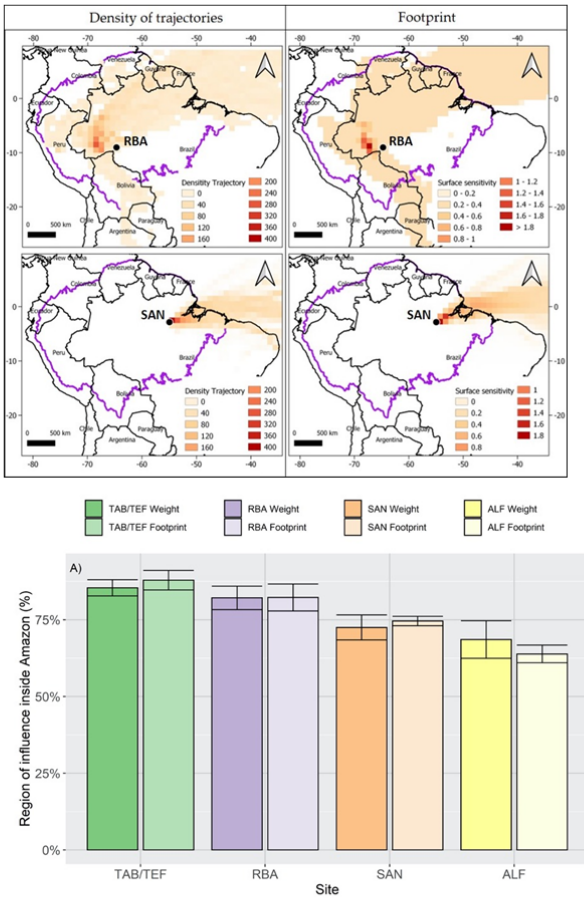


Figure – Left: Comparison of the spatial distribution of the regions of influence provided by the density of trajectories from the method described in the text (left) and the surface sensitivity provided by FLEXPART footprints (right). Both datasets are on their original scales. Right: Difference between the region of influence inside Amazon by footprint used by Alden et al. (2016) and the proposed. Error bar is the standard error at a 95% significance level. Source: Cassol et al. (2020), with permission.

We also tested the influence of the height in the Region of Influence. The reason to choose 3.5 km height was the frequency of biomass burning plumes.

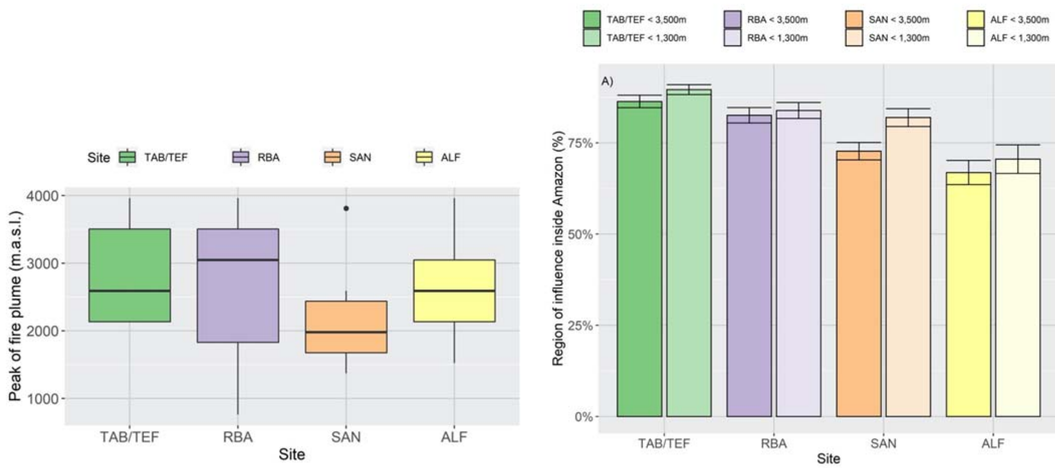


Figure - Left- Distribution of the peak of fire plume in meters above sea level (m asl) at each study site. Right: Source: Difference between the region of influence with binned back-trajectories up to planetary boundary layer (300 - 1,300 m) and up to 3,500 m at each study site in the Amazon mask. Error bars are the standard error at a 95% significance level. Cassol et al. (2020), with permission.

Dear Scott Denning we appreciate your very valuable suggestions, requests for modifications, criticism and questions that help us in produce a better version, more clear of our paper.

Scott Denning

Referee #2 (Remarks to the Author):

A: Summary of the key results

This paper uses a unique (not reported elsewhere) aircraft profiling dataset over Amazonia, covering nearly a decade from 2010. It shows that the eastern, and especially south eastern, Amazon has become a net C source to the atmosphere, driven by a combination of human disturbance and climate change.

B: Originality and significance

While not surprising, the quality of the dataset and narrowness of the error bars make the study results highly significant and novel. Being able to pick apart the fire emissions from the forest net biome exchange figures is a really useful factor of this method, and the long time series with consistent data enables trends to be seen.

The specific climate findings (lines 103-109), while not the focus of the data collected in this study, also struck me as fascinating.

While I have seen some of these data before, I have not seen the results but in such a striking way, nor such a strong impact highlighted: they show the 3 dry season months have experienced warming of 1.37 degrees over Amazonia, with rainfall in these months down 17% - compared to 1 degree of warming and no change in precipitation for the whole year. This strong seasonal stress if vital to understanding the findings in the paper, and has wide consequences. It seems the Amazon dry season is experiencing worse climate change that the average, with similar devastating consequences as the polar regions are experiencing. And the highlighting of this being more extreme in the SE site – 2.5 degrees higher in the dry season, is shocking – unsurprising the forest Net Biome Exchange (excluding fire) is positive here, driven by climate change.

C: Data & methodology: validity of approach, quality of data, quality of presentation

The sample size (600 flights from 4 sites) is very large, and sufficient to draw the conclusions the authors offer. The Regions of Influence calculations appear robust, though this is not my specific area of expertise so I cannot comment on the appropriateness of the method use.

In general I find the methods logical and well explained, and the approach robust. Some figures are a bit difficult to read, or don't show the data to its full potential impact – e.g. Fig. 2 where the text is small and the panels hard to compare.

Authors answer: That is true, thank you in advise us. The Figure 2 (now Figure 4) was modified with text in bigger size.

D: Appropriate use of statistics and treatment of uncertainties

Uncertainties are well covered – the Monte Carlo approach followed appears well justified, and results are mostly presented with confidence intervals that are reasonably calculated.

Authors answer: Due to the questions from reviewer 1, we revised all uncertainties and decided to a less conservative approach to consider the uncertainties for the time trajectory and background. This procedure was adopted to compensate, partially, the losses due to atmospheric convection during the air masses transect crossing Amazonia.

Lines 182-184, using a simple multivariate linear correlation to explain the important change in NBE from sink to source is a bit unsatisfying, if not strictly wrong.

The number of data points may not allow for anything more significant, but it would be good to explain which of these (temperature or water storage) was stronger, and if there was any interaction in the two. And an explanation should be included as for why GRACE data was used rather than VPD or precipitation

Authors answer: It is important to highlight each year mean is composed by 12 months that is composed by 1 or 2 vertical profiles and each vertical profile comes from 12 flasks. Since the mean by year vertical profiles (VP) is around 20 VP, each year mean is composed by 240 samples. Then 9 years is composed by 2160 samples, which can be considered a robust population. We worked with the annual mean to avoid the seasonality that could shows a false correlation. Then the conclusion that each year the NBE is higher is supported by statistical analysis. The Extended Data Figure 5a shows the confidence's interval of NBE vs year and we add 2 more methods to check the linear correlation.

Pearson's correlation r: 0.68 and p: 0.046

Spearman's correlation r: 0.72 and p: 0.037

Kendall's correlation r: 0.56 and p: 0.045

Precipitation studies are presented in Figure 4 (the 4 sites 40 years precipitation upper panel), at Extended Data Figure 8 (the 4 sites 40 years precipitation upper panel) and at Extended Data Figure 3 and 4. GRACE (Gravity Recovery & Climate Experiment), which reflects land water mass anomaly given as equivalent water thickness. Comparing with precipitation, we observed that GRACE reflects the precipitation from 2 months before, (amount of water in soil) the correlation between precipitation with 2 months "lack" shows r: 0.88. The water availability in the soil is more important for the forest than the amount of water in the atmosphere (VPD). Initially we study precipitation and GRACE to observe the water availability correlation with carbon fluxes. At this step we calculated VPD and tested if the correlation could increase. Paired correlations with all parameters don't explain NBE, but a composed correlation with Temperature and water availability in the soil explain better the NBE increase along the years. We test also if have a correlation between Temperature and GRACE, but it can be observed in the table below that they are not correlated. We present a table with separated correlations of NBE with the parameters and combinations between them. As requested by the reviewer we also present all results.

			valid < 0.05
Parameters	Versus parameters	r	p-value
NBE	**VPD**	**0.30**	**0.002**
NBE	**T, GRACE**	**0.88**	**0.011**
NBE	**T, GRACE, VPD**	**0.88**	**0.040**
NBE	**year**	**0.68**	**0.045**
T	Precipitation (P)	0.60	0.089
NBE	GRACE	0.60	0.091
NBE	T, P, GRACE, VPD	0.90	0.094
T	GRACE	0.48	0.187
GRACE	VPD	0.48	0.190
GRACE	Precipitation (P)	0.47	0.204
NBE	T, VPD	0.59	0.270
NBE	T	0.28	0.470
VPD	Precipitation (P)	0.17	0.668
NBE	Precipitation (P)	0.13	0.730
T	VPD	0.06	0.886

Linear Regression										
Dependent variable		NBE								
Independent variables		Temperatura, GRACE								
N		9								
Regression Statistics										
R		0.86942	R-Squared		0.75589	Adjusted R-Squared		0.67452		
MSE		0.00685	S		0.08277	MAPE		75.84372		
Durbin-Watson (DW)		2.22364	Log likelihood		11.47884					
Akaike inf. criterion (AIC)		-1.88419	AICc		-1.66196					
Schwarz criterion (BIC)		-1.81844	Hannan-Quinn criterion (H		-2.02606					
PRESS		0.08964	PRESS RMSE		0.0998	Predicted R-Squared		0.46769		
NBE = 9.70279 - 0.36715 * Temperatura - 1.94970 * GRACE										
ANOVA										
	d.f.	SS	MS	F	p-value					
Regression	2	0.1273	0.06365	9.28959	0.01455					
Residual	6	0.04111	0.00685							
Total	8	0.16841								
	Coefficients	Std Err	LCL	UCL	t Stat	p-value	H0 (5%)	VIF	TOL	Beta
Intercept	9.70279	3.05462	2.2284	17.17717	3.17643	0.01916	Rejected			
Temperatura	-0.36715	0.11616	-0.65138	-0.08292	-3.16075	0.01955	Rejected	1.29971	0.7694	-0.72683
GRACE	-1.9497	0.47688	-3.11658	-0.78282	-4.08847	0.00644	Rejected	1.29971	0.7694	-0.94016
T (5%)	2.44691									
LCL - Lower limit of the 95% confidence interval										
UCL - Upper limit of the 95% confidence interval										

I'd like the authors to present the data to check they've not just chosen the variable with the highest correlation, but then not explained it's lower for the others.

Table ALF Annual mean fluxes and precipitation, temperature, GRACE, burned area and EVI

				Precipitation (mm)					Temperature (°C)							
ALF	NBE C Flux (gC m^{-2} d^{-1})	Fire C Flux (gC m^{-2} d^{-1})	Flux (gC m^{-2} d^{-1})	Annual	Wet season (NDJFMAM)	Wet peak (JFM)	Dry season (JJASO)	Dry season (ASO)	Annual	Wet season (NDJFMAM)	Wet peak (JFM)	Dry season (JJASO)	Dry season (ASO)	GRACE (m)	Burned Area (km^2)	EVI
2010	-0.08	0.26	0.19	1968	1555	753	334	194	26.8	26.5	26.4	27.5	28.0	-0.0498	567.97	0.488
2011	0.04	0.16	0.20	2143	1862	1044	315	284	26.2	25.4	25.1	27.3	27.9	-0.0120	136.47	0.496
2012	0.11	0.22	0.32	1984	1695	1032	271	199	26.1	25.5	25.1	26.8	27.5	-0.0229	425.36	0.487
2013	0.22	0.13	0.36	2073	1779	996	343	240	26.2	25.7	25.6	26.8	27.3	-0.0002	122.61	0.500
2014	-0.06	0.17	0.11	1929	1687	982	308	253	26.4	25.8	25.6	27.2	27.3	0.0231	181.70	0.502
2015	0.08	0.22	0.30	1561	1486	836	207	111	26.7	26.0	25.5	27.7	28.3	-0.0915	442.27	0.482
2016	0.34	0.15	0.48	1870	1193	776	357	297	26.6	26.5	26.4	26.9	27.2	-0.2151	230.79	0.489
2017	0.13	0.27	0.41	1895	1766	1003	191	164	26.6	25.8	25.4	27.7	28.4	-0.0674	713.92	0.480
2018	0.28	0.13	0.41	2009	1645	794	332	253	26.0	25.6	25.3	26.8	27.1	-0.0430	106.97	0.490

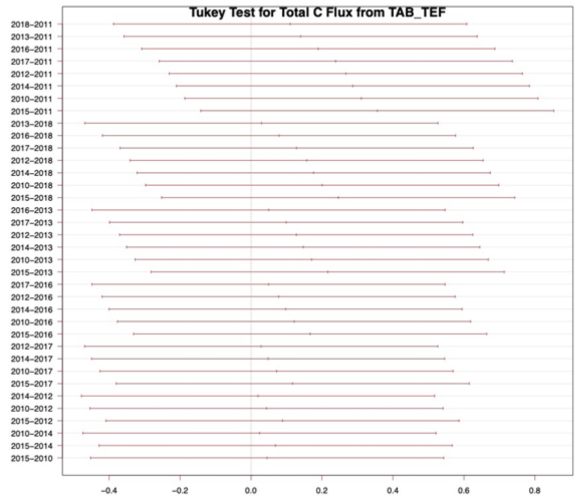
We're left to take the author's word that TAB (2010-12) and TEF (2013-18) are similar enough to be analyzed as one trend. This looks reasonable from the map shown in Figure 1, and the evidence of no discontinuity in Extended Data figures 6 and 7c – but at the very least a line Extended Data Fig.6 showing where the two sites switched over would be useful. And I think it would be useful to have a specific figure/table covering this issue, giving the average ROI (Region of Influence) covered by each, %historical deforestation of each, and how these change with the time of year – to prove this is acceptable.

Authors answer: It was included a time series Figure for TAB (2010-2012) and TEF (2013-2018), with the background stations (RPB, ASC and CPT) and the annual deforestation at Supplementary Information Figure 2.

We also done a Tukey test:

The Tukey HSD test (Honestly Significant Difference) is a statistical tool to determine whether the relationship between two groups of variables is statistically significant. In other words, if there is a chance that a deviation from an observed numerical value is casually related to a deviation from another observed value, it is a practical way of testing hypotheses.

The Tukey test was applied to the Total Carbon Flux data for the years 2010 to 2012 for (TAB) Tabatinga and 2013 to 2018 for (TEF) Tefé. The figure below is calculated with a 95% confidence interval. When the horizontal data pair is cut by the vertical dotted line that passes through zero, you can assume that the compared pair has no statistically significant difference, which corroborates the similarity of the data between the two studied sites. As an example, we have that the biggest difference observed occurs between the years 2011 and 2015, but still considered similar.



It is a shame there is so much missing data, but the length of the time series and the 'MissForest' approach taken to gap fill seem to be okay. I don't think these gaps, if filled with real data, would change any of the conclusions.

Authors answer: This is the hardest place in Amazon to sample. The logistical difficulties are very hard, and also for the airplanes companies to maintain open and operational it is also challenge. At August 2012 the airplane, based at Eirunepé, Amazonas state, crashed with a motorcycle and the driver died. The police take the airplane to investigate, and we needed to move to another site. We founded a company in Tefé, Amazonas, not ideal, but feasible. It is really difficult the logistic in send the samplers to take 2 samples per months. Sometime one direction of shipment was done part by airplane, truck, boat and truck again (take 20 days to total itinerary). 2 years ago, the unique airplane had also an accident and we paid more than doble price because the fly started from Manaus, during more than one year for don't stop the samples, but was done at this time only one vertical profiles per month. So, there are many logistic difficulties to maintain 2 samples per months at this remote site in Amazonia.

E: Conclusions: robustness, validity, reliability

I have no doubt that the conclusions are robust, valid and reliable. The conclusions come directly from the data collected, and the implications are self-evident. While alluded to a little at the end, I do have a slight worry that the NBE figures are treated as comparable to the fluxes due to intact forest, e.g. as compared to the Hubau et al. (2020) paper in the final paragraph of the main text. These fluxes do not just contain intact forest – they contain many areas of Human Modified Forest, including logging concessions actively undergoing logging and recovering from past logging, and forests undergoing recovery from other forms of disturbance (e.g. last year's fires) – plus non-forest vegetation will be contributing to these figures too. This is not a problem with the data or analysis – the facts and figures stand – but I would like more notes of caution from the authors around these areas, especially to encourage people to be careful of comparisons.

<u>Authors answer</u>: Thank you in point this valuable comment. We change the discussion about the results and highlighted that our NBE includes all fluxes natural and anthropogenic non-fire emissions and absorptions. This part of results analysis and comparisons are in lines 172-188, and the comment more specifically at lines 185-188.

F: Suggested improvements:

Small point, but neither reference 7 or 8 give information on 'below-ground organic carbon'. 7 does include a belowground biomass layer, but adding that to the AGB gives a total value for Amazonia nearer 100 Pg than 200 Pg C. Their number sounds about right for a figure including soil carbon, but they need to fix the references.

<u>Authors answer</u>: Thank you in point this mistake. The references were changed to a study that include biomass above and below ground. The reference now is the 7 Malhi et al, 2006.

Lines 126-127 – it is strange to give so many (useful and interesting) facts and figures about temperature rises in the dry season for each site – but not to give figures for precipitation and VPD changes, instead just referring to the literature. I think it would be helpful for the reader to have these listed, or summary figures could be added to Fig. 2 to assist with comparing between panels? I know these might duplicate those in Fig. 4, so maybe this isn't necessary – but Fig. 4 only has the ASO figures, not those for the whole year. VPD seems important too, was it not possible to quantify changes in this over the sites?

<u>Authors answer</u>: Precipitation studies are presented in Figure 4 (the 4 sites 40 years precipitation upper panel), at Extended Data Figure 8 (the 4 sites 40 years precipitation upper panel) and at Extended Data Figure 3 and 4 (9 years study). At this step we calculated VPD (only for the 9 years study) and tested if the correlation could increase, since we didn't observe a better correlation than using precipitation or GRACE (water availability to the forest), we didn't modify this part of the paper.

The units throughout the results are given in gC m-2 d-1. This is a perfectly acceptable unit given what is measured, but is hard to compare to the fluxes given in Pg carbon, e.g. in the first sentence of the main paper, and in the discussion of carbon sinks/sources here and elsewhere (e.g. line 240). I wonder if units of Pg C yr-1 could be easier to compare, though they would need to be standardized by area somehow so might get unwieldy.

Authors answer: We presented the carbon flux, because you can transform in PgC y^{-1} using a different Amazon area. We adopted Amazon region (Supplementary Information Figure 3). To integrate the carbon Balance for the whole Amazonia we separated in 3 different regions (Extended Data Figure 6a), like described at lines 753-761.

Line 528-529, I would have liked to have seen information as to how the Burned Area product was resampled to 1x1 degree resolution. The original 500m data are binary (the whole pixel is burned on a particular calendar day or not). Was the ultimate 1 degree product, as I assume, the proportion of that 1 degree pixel that burned within a particular time period (e.g. 3-month period)? I assume this has been done properly, but it should be clearer in the text.

Authors answer: Thanks for addressing your concern. Yes, to resampled to 1x1° resolution, we used the fraction of the grid cell that was burned and then transformed to total area and period for each month (lines 675-676). We have included this piece of information in the revised manuscript: "*The BA product was resampled to 1x1° spatial resolution using the fraction of area burned in that grid cell and summed for each quarter in the IDL/ENVI®*" (lines 680-682)

Lines 531-542, would be good to have justification for why EVI was used, rather than say NDVI, and why given this is about (according to this paragraph) fPAR, the explanation then talks about 'greater biomass' – fPAR and biomass are not well correlated (nor are EVI and biomass beyond a certain quite low level). I don't think anything has been done incorrectly here, but I think a bit more thought should be shown as to what EVI is showing (definitely something about leaves more than biomass – the seasonal trends seen in EVI in the paper are not seasonal biomass trends!) and why it is chosen (lower artefacts than NDVI typically through inclusion of blue band).

Authors answer: Thanks for your comment. The EVI presents adjustment in the signal/noise by including the blue band and atmospheric corrections that minimize the influences of the aerosols. The MODIS MAIAC data is adjusted to solar-nadir capturing, minimizing the artefacts provided by multiangle in a large field-of-view of the sensor. We have changed the sentence to provide a justification of chosen EVI in the analysis: "*In general, high positive values show a higher proportion of fPAR, and therefore, greater vegetation greenness (vegetation vigor). EVI can also reveal the seasonality of different vegetation types, where tree-individuals partly lose leaves during the dry season, become drier, and reducing the index value*. *Contrariwise to NDVI, EVI includes a blue band that minimizes the influence of aerosols and other adjustments to improve signal/noise ratio*" (lines 684-692).

There are climate models that make predictions about carbon fluxes (including and without fire) at a sub-basin scale. The authors will certainly be aware of these, with several involved in intercomparison papers – it would be interesting for the reader to know if in general terms the eastern vs western Amazon are behaving as models would expect given this level of climate change an atmospheric CO2 concentration, or whether these are outwith the expectations of the majority of models. This may be too big a separate piece of analysis, but if relatively trivial to do and explain I think it would add a useful 'implication' paragraph in the main text of the paper, either validating the models we are using, or suggesting they are not performing well in a critical region.

Authors answer: We appreciate the suggestions in include the comparison with others approaches. The discussion about the Amazon Carbon Balance and previous studies for comparison now is presented from line 173 to 188. It is clear the discrepancy from bottom-up approach since forest plots at mature forests and the top-down studies like satellites and vertical profiles, that capture all the process: sinks and sources. More comparison was presented at Supplementary Information Table 2.

G: References: appropriate credit to previous work?

Relevant papers appear to have been cited.

H: Clarity and context: lucidity of abstract/summary, appropriateness of abstract, introduction and conclusions

I believe the paper is well written overall, with some smaller problems as highlighted above.

The authors thank you for the careful review and the valuable questions and suggestions that improved the paper.

Referee #3 (Remarks to the Author):

The manuscript presents an evaluation of the Amazon carbon balance, using air column measurements at 4 sites with a basin of observation that covers substantial portions of the Eastern and Central Amazon. The key observations are 1) that the Amazon has been a net carbon source from 2000-2018 and 2) that fire and deforestation have made the Eastern Amazon more of a source than the western Amazon. The dataset is rich and seemingly definitive, and it is on the merits of the empirical work that the manuscript merits consideration for publication in Nature.

Personally, I see three sites acting as a sink, and only one site in the SouthEast acting as a source (and not the entire Eastern Amazon as currently framed). Overall, the data and trends seem quite logical and consistent with the published literature to me, although I would reframe the summary of the results.

On that subject, I'm surprised to see so little context w.r.t. what is known about Amazon-wide carbon fluxes especially about fire, since the authors argue that fire is driving source vs. sink outcomes. Fire in the Amazon is quite well studied but mostly neglected as a point of discussion in the manuscript. I can find only one publication (#10) that deals with fire fluxes in the entire paper, which is quite thin given its central importance as a driver of basin-wide carbon dynamics (in all regions, frankly).

Authors answer: Thanks for addressing your concern in this topic. We improve the discussion in the paper (lines 174-184) and cited more studies for comparison. We also add a Supplementary Information Table 2.

Overall, I found the manuscript somewhat difficult to read. They rely heavily on site acronyms in the results section, which requires a reader to go back and memorize which is which, so this really impedes clear communication. The results and discussion should really be reorganized if the target is a short format journal – instead of presenting the drivers and carbon results separately, these should be integrated into a region-by-region combined results and interpretation. Then end with a substantial section evaluating the effects of different drivers on spatial and temporal trends in carbon fluxes across the basin.

Authors answer: Thanks for addressing this difficulty. We change the code for each site to include at name site the region from geographical position. Now SAN is SAN-NE (northeast)… as detailed at lines 90-94. The paper structure was also modified for a better understand.

The new structure now is:

a) Vertical profile data and studies based in CO_2 concentrations.
b) Flux results with Column budget technique and in the last part the C balance

c) Climate analysis for the last 40 years (temperature and precipitation) and deforestation for each region of influence.
d) Comparison east versus west in Amazon regions
e) Hypotheses about the link between the carbon fluxes and balance results and the human impact at it regions with deforestation and changes at dry season.

Fig 1 but also a larger methodological point: it's convenient that the site areas match up exactly with the Amazon mask. Obviously, this is not how it actually works – ALF for instance might include a significant non-forest component in the atmospheric signal, so too RBA (whereas San and TEF presumably include a lot of ocean in the signal which may also be an issue?). In the case of RBA, some substantial component of the fire signal may be coming from the Cerrado, so that may be misattributed here. This needs to be dealt with directly, at least in the figure but more likely in the method overall.

Authors answer: Thanks for addressing your concern in this topic. We include in the text the lines from 137 to 143 to make clear the possible influence of Cerrado at east sites. Since we subtract background a priori we are removing the ocean contribution. This is an usual procedure of the used method "Column Budget Technique. The background method description is presented at lines 542-550 and in a new Supplementary Information Figure 5.

Line 37: Even a few words about what the framework, instead of just calling it "simple"

Authors answer: We rephase the lines 37 - 39

Line 40: remove "relatively pristine" – unnecessary value judgement

Authors answer: We rephase "less human impacted" at lines 41-42.

Line 43-44: something funny about the sentence construction here. A flux can't be a source, although a region can be. Rephrase.

Authors answer: We rephase in the lines 43-47.

Line 53: "around 2200 mm per year" This is a weird number to put here, since it varies substantially across the basin. Either remove entirely, or add "on average, across the basin, rainfall around 2200 mm per year"

Authors answer: We rephase in the lines 53-54.

Line 112: It's quite an empty statement to say that 2 out of only 4 stand out. Are you defining the mean based on the other two?!

Authors answer: We restructured completely the manuscript. The topic "*Climate trends and human impact"* now is in the lines from 186 to 236. The comments about east sites are until 221 and the west-central sites in the following paragraphs.

Line 155: On a semantic level, it's counterintuitive to me that fire is not a part of total flux (although obviously it isn't a part of NPP). This may be standard terminology, although I have not encountered it, but at a minimum I would be sure to include definitions of these in the captions for figures where fluxes are presented. It's a real schlep to sift through the text to find a definition in order to interpret a figure.

Authors answer: Thank you in point this important information. This is a kind of different views from communities that work more with bottom-up data and the community using atmospheric data. We clarified our definitions in the lines from 121-124. Since we measure CO_2 in the atmosphere, total for us is correlated to the CO_2 concentrations founded in the atmosphere. It is the resultant of all processes: natural and anthropogenic. We have only one tool to distinguish the CO2 coming from biomass burning using CO as a tracer. We clarify also the signal that represents carbon source or sink (lines 124 and 247).

The authors thank you for the careful review and the valuable questions and suggestions that improved the paper.

Reviewer Reports on the First Revision:

Ref #1

This is one of the most ambitious and potentially impactful manuscripts I've ever reviewed. The authors report results of nearly 600 vertical profiles of trace gas measurements made from small aircraft across the Amazon forest over 9 years. It's an unprecedented data set and shows distinct and provocative patterns in space and time. The authors highlight the difference between small carbon losses in the eastern Amazon and small carbon gains in the western Amazon, especially due to fires in the dry season. Using climate and land use data, they propose a mechanism to explain these variations as a result of seasonal warming and drying as well as deforestation. The results will impact our understanding of the carbon balance of tropical forests, the vulnerability of their carbon stocks, the dependence of terrestrial carbon sinks on climate change, and the future trajectory of carbon-climate feedback in the Earth system.

The authors have done a remarkable job of restructuring the manuscript, documenting their methods, and analyzing/quantifying the sources of uncertainty in this revised manuscript. I am really gratified to see the improvements, and I think this is a critical and timely contribution to this important field.

I recommend that the manuscript be published in Nature without further delay.

Ref #2

The authors appear to have done an excellent job of responding to the concerns I and the other reviewers had. The fundamental findings of the paper have not changed, but the description of methods and detail in some of the figures is now better.

I remain highly impressed with the dataset they have managed to put together, and the results they produce are important.

I believe this should be published as is.

Ref #3

The manuscript has quite a large number of grammatical errors; I suggest careful proofreading either by the authors, an outside proofreader, or by proofreaders associated with Nature, at the discretion of the editor.

On the one hand, I appreciate the addition of geographical information to site abbreviations, but on the whole this addresses the letter of my suggestion rather than its spirit. The manuscript remains difficult to read in part because of an excessively large number of acronyms. I would also note that this geographical information has been added to abbreviations in an incomplete fashion, and should the manuscript proceed to the next stage, I recommend that the authors should carefully review all acronyms to ensure that they are complete/correct.