Response to Anonymous Referee #1

This article presents original estimates of evapotranspiration (ET) derived via water balance in the Amazon and several of its subbasins. The authors use these estimates to examine spatial variability in ET within the Amazon, seasonal and interannual variability in ET at the subbasin and Amazon-scale, and the dominant drivers of ET variability in the region. The authors additionally compare their findings to existing ET data from an extensive suite of remotely-sensed ET products, site-scale ET from flux towers, re-analysis models, and CMIP GCMs to evaluate how well they capture the dynamics of Amazonian ET.

I commend the authors for writing a clear and detailed article that will be of interest to both climate modeling and tropical ecohydrology audiences. Their references to prior literature are thorough and concisely summarized, their findings are presented in well- designed figures, and the details included in the supplementary information will allow readers to build a deep understanding of their analysis. I am also happy to see a clear explanation of error propagation in the paper, which is important for any study using a water balance approach as a benchmark for more complex models.

Answer: We would like to sincerely thank you for taking the time to carefully read and review our work. We have revised the manuscript according to your suggestions, and feel that is now a stronger paper as a result. We respond to each of your comments in the text below.

I do not think any major revisions or additional analyses are needed, but I have some comments and questions listed below that might improve the paper:

L24: The abstract reports a "strong seasonal cycle in basin-mean ET controlled by net incoming radiation...". Here and elsewhere throughout the paper I would consider slightly softening the language, e.g. by saying that basin-wide ET is "primarily controlled by radiation" or "highly correlated with radiation" or something similar. My concern is that while the current statement is consistent with your findings, it is very general and may mislead unfamiliar readers into neglecting other important dynamics behind Amazonian ET. As your results and many other papers have found (Maeda et al. 2017 in ESD comes to mind), vegetation dynamics and water availability (via rainfall as well as terrestrial availability) appear to modulate the ET signal in certain regions during certain times of the year. You have demonstrated quite well that, when aggregated across the entire basin, ET and Rnet have strikingly similar seasonal cycles. But I would be cautious about unintentionally implying radiation is the sole control on ET's seasonality when the full story is more complicated.

Answer: This is a good point, and we have now amended the text to be more careful in our language when describing the controls on ET.

L55: The comma after ET seems unnecessary to me.

Answer: We have removed the comma.

L199: There appears to be some formatting mistake with the uncertainty term here.

Answer: We have corrected the typo.

L227: If pasture is the dominant regional land use, why exclude these tower sites? Is pasture just not common enough across your subbasins to be reflected in your tower analysis?

Answer: We apologise for miswording this sentence. The pasture towers we excluded were in areas where the dominant land cover was forest, rather than pasture, and thus the towers were not representative of the surrounding land cover. We have corrected the text.

L234: Suggest removing the "to" in "near to Manaus"

Answer: Corrected.

L249: It's worth including the version of MOD16A2 that you used, presumably version 6. I would say the same thing for P-LSH and CHIRPS if there are publicized version numbers associated with the data (I think CHIRPS is on version 2 these days).

Answer: Thanks for this suggestion. We have added the version numbers for MOD16A2 and CHIRPS to the text. To our understanding, there is no published version number associated with the P-LSH dataset.

L279: Why re-grid to 1x1 degree pixels for the visualizations? To reduce noise? I would mention somewhere what the native resolutions of the CMIP models are before resampling.

Answer: The CMIP models vary from model to model in their resolution, and in order to calculate a multi-model mean it was necessary to harmonise the models to a consistent resolution. We selected 1° x 1°, which is towards the finer end of the model resolutions, as this enabled us to extract data from each Amazon sub-basin with more accuracy than with a coarser grid. We have added columns to Tables S3 & S4 giving the native resolution of each model and added a line to the text emphasising that though not all models simulate the level of detail provided by 1° x 1°, choosing this resolution enabled us to extract data from each Amazon sub-basin with more accuracy than using a coarser grid.

L295: How many of these gaps were there? If I'm understanding this correctly, each time series is only 11 datapoints, so filling any of those datapoints with climatological means could obscure trends quite a bit. Worth mentioning somewhere in the text or Supplementary Info how much gap-filling was necessary for each basin's ET time series.

Answer: Sorry this sentence was included in error. We calculated interannual trends for the Amazon basin only and had originally used a gap-filling approach for the catchment-balance series, but on reflection decided to exclude years with fewer than 10 months of data, or any years with a missing month for the JFM and JAS time series. Therefore, years 2017 and 2018 were removed from the annual time series (Fig. 8a), and the year 2017 from the JAS time series (Fig. 8c). The trends calculated for 2003–2013 were unaffected by data missing from these years. We have amended the text to clarify and correct the description of the methods.

L304: On first read I was a bit confused by the term "corresponding sources." In fact this whole paragraph could use some small tweaks for clarity. This reordering of the sentences (and complete removal of the "corresponding sources" sentence) sounds better to me, but you have already demonstrated that you are a capable writer so I will trust you to make whatever changes you see fit: ". . .Bins with fewer than five data points were excluded from the analysis. Satellite-based ET estimates were binned according to precipitation from CHIRPS, radiation from

CLARA-A1 (Karlsson et al., 2013) and LAI from the MODIS MOD15A2H product (Myneni et al., 2015), each re-gridded to 0.25◦ x 0.25◦, while reanalysis and model ET were compared with reanalysis and model variables, respectively. For ERA5, we used the 'high vegetation' LAI field since the Amazon is predominantly covered with tropical forest, though repeating the analysis with 'low vegetation' LAI made little difference to the results. Note that the satellite-based MODIS ET. . ."

Answer: Many thanks for this helpful suggestion to improve the clarity of the text. We have amended the paragraph in line with your suggestions and agree that it now reads more clearly.

L306: I would appreciate some very brief discussion of how trustworthy CLARA and MOD15A2 can be considered in this region, if anything is known. For instance, does MOD15A2 suffer from the sun-sensor geometry issues known to affect MODIS in the Amazon? I don't know much about the CLARA-A1 dataset but perhaps there is some validation study available, or at least you could explain why you chose it over other radiation datasets.

Answer: This is an important point, and we have added some discussion on the validity of these products over the Amazon to section 2.6:

"MODIS LAI has been shown to perform relatively well against ground-based LAI measurements ($R2=0.7–0.77$), though uncertainty over the validity of high LAI values (>4 m^2 m^{-2}), such as occur over the Amazon, is larger due there being few ground measurements and the satellite reflectance signal reaching saturation over dense canopies (Yan et al., 2016a). Furthermore, the satellite-based MODIS ET product incorporates MODIS LAI (Table 1), and thus these datasets are not fully independent from one another. The CLARA-A1 radiation is independent from the ET datasets evaluated in this study and estimated to have an accuracy of ≤10 W m^{-2}, though few surface measurements were available over South America, and none in the Amazon region (Karlsson et al., 2013). Thus, there is some uncertainty in the accuracy of these satellite products over the Amazon that must be considered when interpreting the results."

L310: Was the Amazon's hydroclimate during 1994-2004 broadly similar to 2003- 2013? I'm not immediately familiar with the Amazon's recent climate trends, but if one period had worse droughts or wetter wet years than the other, mean ET may show a response to that. Somewhere in the paper, one or two lines covering this question might be a good addition.

Answer: You are right that there has been a change in the Amazon hydroclimate between these two periods, with the Amazon hydrological cycle becoming more intense and basin-mean P increasing (Gloor et al., 2013). Therefore, we might expect to see some differences between CMIP5 results and results from other data sources. However, we found CMIP5 ET to be largely consistent with ET from CMIP6, despite the different time periods analysed, suggesting differences due to time period were smaller than differences between the models and other types of ET data. We have added a comment on this to section 2.6.

Figure 3: Is there some kind of interpolation or smoothing being used here? My understanding was that the CMIP models were resampled to 1-degree grid cells for visual- ization, which seems much coarser than the data in panels g and h. I don't necessarily take issue with interpolating for the presentation of this figure, but it may provide a false sense of spatial detail and should be explicitly stated somewhere.

Answer: You are right, we used the filled-contour plot function to map the data, with contours at 25-mm intervals between 1000 and 1500 mm $year^{-1}$. We have now added this information to the figure caption for transparency.

L354: I don't know if I'd say the tower gradient is "similar" given so few datapoints, and the two northeastern towers featuring such different values. Perhaps just say they appear to display an east-west gradient. Some explanation of why the two very close towers feature quite different mean ETs would be welcome as well – is it a land cover difference?

Answer: We have amended the text to make a more careful comment on the possible gradient suggested by the flux tower data. We have also added a statement stating that the observable difference in mean annual ET between the two nearby towers in the northeast amazon is likely due to these being on different land-cover types (primary forest and selectively logged forest).

L377: As I wrote regarding the abstract, I would be careful with statements like "water availability is not a limiting factor controlling the spatial distribution of ET over the Amazon." It seems to me that P may well limit ET in some catchments (e.g. Madeira, which has relatively low ET and P but relatively high RDN in your plots. Maeda et al. (2017) characterized Madeira and other basins as water-limited for at least part of the year). I think you could justifiably conclude "water availability does not consistently limit ET in all regions of the Amazon" from Figure 4, or that "water availability does not limit ET as consistently across Amazonian subbasins as radiation." But as it is now, I think the line is potentially misleading.

Answer: Thanks for drawing our attention to this ambiguously-worded sentence. We did not mean to give the impression that water-limitation never occurs over the Amazon, rather that spatial variation in ET across the Amazon appeared to be associated with spatial variation in radiation, and not water availability. We have rewritten the sentence in question and made sure to be more careful with our language: "This result tentatively suggests that spatial variation in radiation explains more of the spatial variability in ET across Amazon sub-catchments than other variables." The text of the whole paragraph is copied at the bottom of this document.

Figure 4 (and others): The symbols for Amazon and Tapajos are difficult to distinguish without zooming in. Fortunately you have prepared your figures well enough that zoom- ing in is possible, but I worry that people reading printed versions of the article would still have trouble. Maybe Tapajos could be replaced with an unfilled circle?

Answer: We have amended the symbol for Tapajos in Figures 4, 5, S1 and S7 to a star which is more distinguishable from the circle than the hexagon, and increased the size of the markers.

Figure 7: I am curious why you decided to do this analysis on monthly climatological means rather than just plotting LAI and ET in every individual month you had available. Also, it looks like you are calling P-LSH "LSE-Zhang" in the plot titles; I would change this to be consistent throughout.

Answer: We wanted to separate the controls on seasonal variation in ET (Fig. 7) from the controls on interannual variation in ET (Fig. S10), which is why we only plotted the climatological monthly means in this plot. We have amended the Method to clarify our reasoning. We have corrected the title for the P-LSH panels – thanks for drawing our attention to this error.

Supplemental Information:

Table S2: Please clarify what area these values pertain to.

Answer: We have added this information.

Table S6: Change Zhang to P-LSH?

Answer: Correction made.

Figure S4: In last sentence of the caption, change "was" to "were."

Answer: Correction made.

Figure S7: I would love to see this figure replicated for RDN and LAI, since so much of your analysis of various ET data sources hinges on the comparisons to their respective RDN and LAI datasets. After reading your article I was left wondering how spatially variable these drivers are in the reanalysis and GCMs.

Answer: This is a very nice idea – thank you for the suggestion. We have expanded the figure to include maps showing climatological annual mean radiation and leaf area index, in addition to precipitation. These maps clearly show that radiation is higher in the models than in the satellite or reanalysis products, particularly in the eastern Amazon. Satellite LAI also shows lower variability than other products, likely due to signal saturation. We have added this to the discussion.

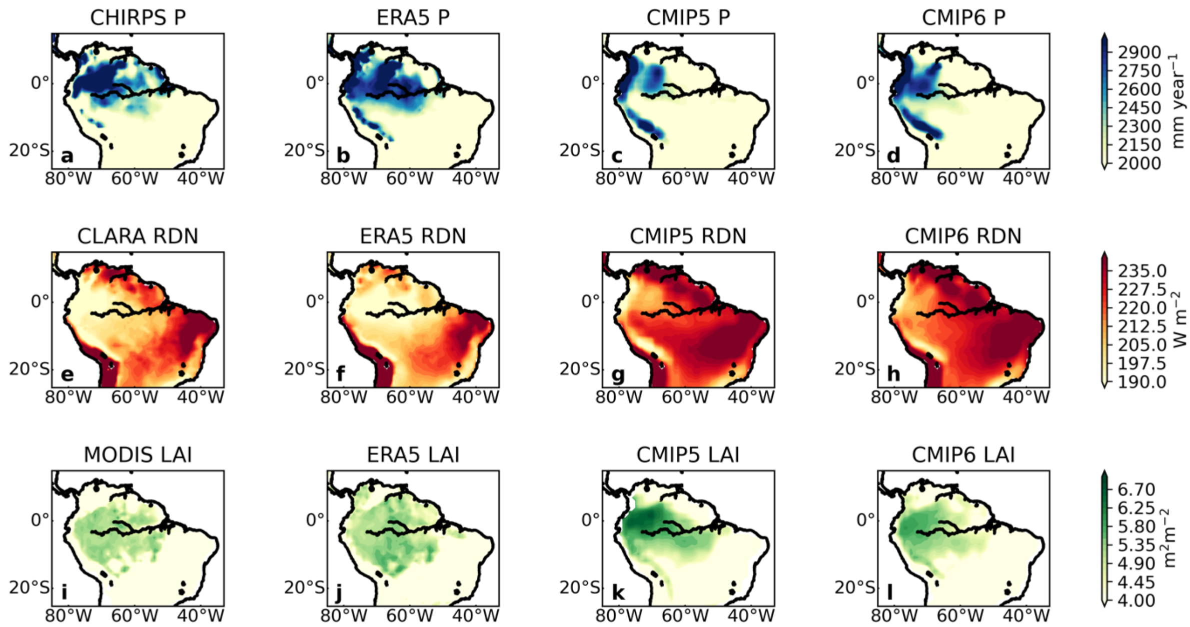

Figure R1 – Climatological annual precipitation, radiation and leaf area index. Mean annual precipitation (P, a–d), radiation (RDN, e–h) and leaf area index (LAI, i–l) from satellites (column 1), reanalysis (column 2) and climate models (columns 3 & 4) over the period 2003–2013. References for each dataset are provided in the main paper.

Figure S9: If it's not too much trouble, a similar plot showing LAI's seasonal cycle across the basin would be interesting since LAI appears to relate to ET when viewed across subbasins and months. Adding to this figure may make it too crowded, so perhaps just in a new supplemental figure.

Answer: Thank you for this useful suggestion. We have added a figure showing the satellite LAI and catchment-balance ET seasonal cycles side by side (Figure S11) and refer to it in the text.

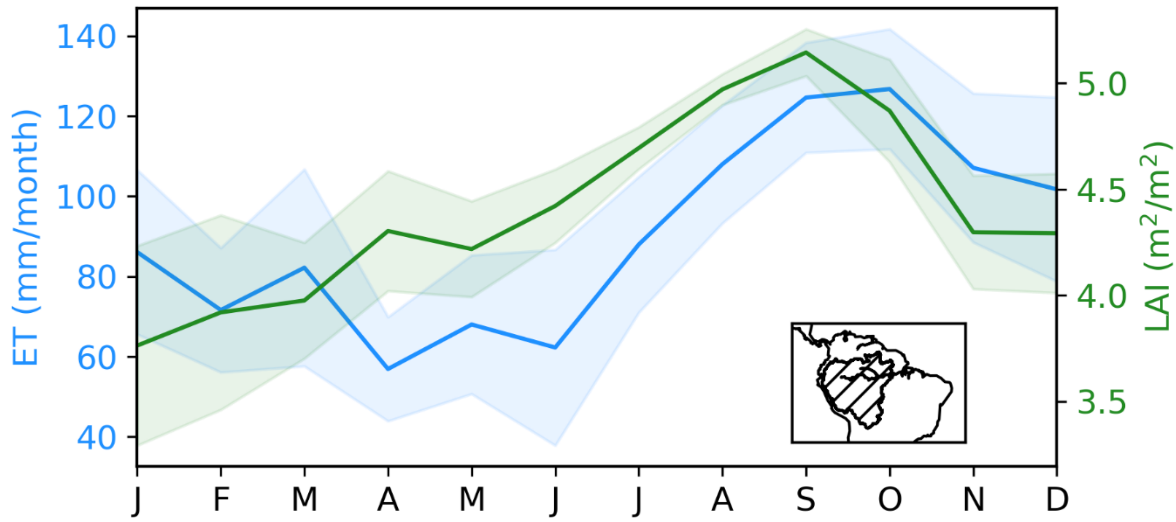

Figure S11 – Seasonal variation in leaf area index over the Amazon. Climatological seasonal cycles in catchment-balance ET (blue) and MODIS MOD15A2H Collection 6 LAI (green) averaged over the Amazon region shown in the inset map (area-weighted mean using data from 2003–2013). Confidence intervals indicate the interannual standard deviation in each month. Note that y-axes do not start at zero.

Figure S10: I don't understand what the colors in the caption are referring to. Where is dark blue? What about red and magenta?

Answer: We have corrected this figure caption (now Fig. S12) and removed the references to colours. Apologies for this error which related to an earlier version of the figure.

Other changes

We noticed that our Amazon LAI values were implausibly low (Amazon mean LAI value of 3.6 m^2/m^2), likely due to inadequate quality control during data processing. We have changed to use a quality-controlled MODIS MOD15A2H Collection 6 LAI dataset provided by Boston University (Amazon mean LAI value of 4.4 m^2/m^2). The main difference to the results arising from this change is that catchment-balance ET is no longer well related to spatial variation in LAI. The new figure and paragraph describing these results are copied below. There were no meaningful changes to any of the rest of the results.

"To understand the drivers of spatial variation in Amazon ET, we compared catchment-scale estimates against catchment-means of precipitation, surface radiation and LAI (Fig. 4). Since there were only eleven data points in the analysis (representing the Amazon and ten sub-catchments), statistical power was relatively low. However, we found spatial variation in catchment-balance ET showed some indication of an influence from radiation ($r=0.38$, $p=0.25$, Fig. 4h), but not precipitation ($r=0.14$, $p=0.68$, Fig. 4a) or LAI ($r=0.06$, $p=0.87$, Fig. 4o). This result tentatively suggests that spatial variation in radiation explains more of the spatial variability in ET across Amazon sub-catchments than other variables. None of the ET products and models analysed captured positive relationships between catchment-mean ET and radiation. ET from ERA5 and the CMIP ensembles instead showed negative associations with radiation (Fig. 4l–n), and, along with GLEAM ET, positive relationships with precipitation (Fig. 4d–g), indicative of water availability influencing spatial variation in ET (Fig. 4d–g). These results confirm that the reanalysis and climate models analysed here struggled to capture spatial patterns in Amazon ET due to misrepresentation of the controlling drivers, specifically the relative importance of precipitation and net radiation. ET from ERA5 and the models also showed positive correlations between LAI and ET (Fig. 4s–u), not seen in the satellite observations. However, it should be noted that satellite LAI was generally lower and showed less spatial variability than other LAI datasets over the Amazon (Fig. S8i–l), likely due to the satellite sensor being insensitive to variation in LAI over areas of dense tropical forest (Myneni et al., 2002, Yan et al., 2016a). This could hamper our ability to accurately assess the extent to which LAI influences spatial variation in ET."

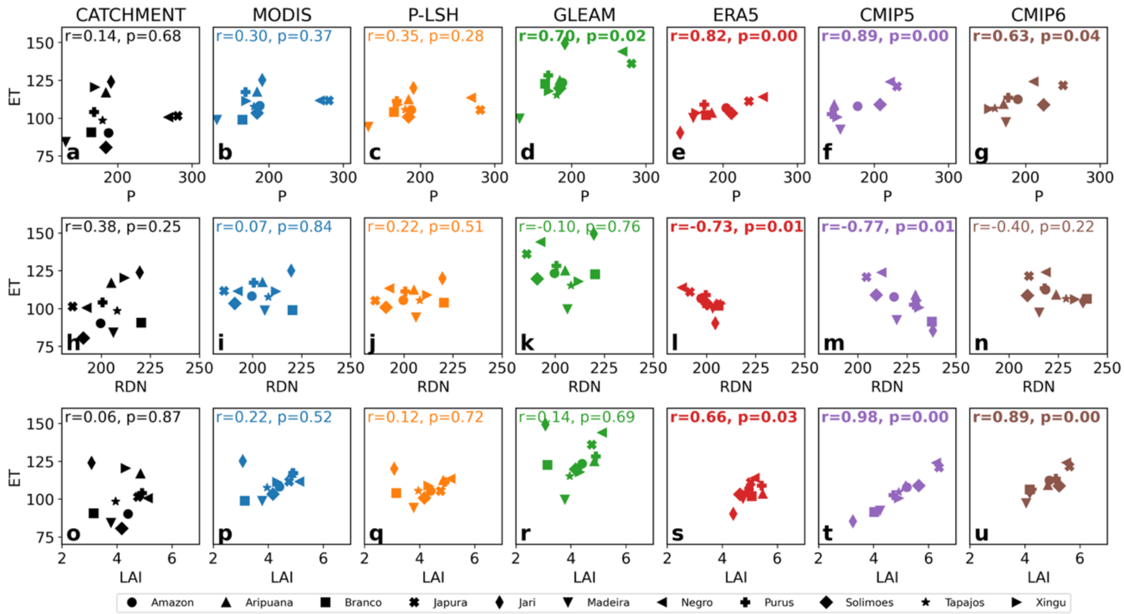

Figure 4 – Controls on spatial variation in Amazon evapotranspiration. Annual mean ET (in mm month^{-1}) for the Amazon and ten sub-catchments (Fig. 1) from catchment-balance, satellites (MODIS, P-LSH, GLEAM), ERA5 reanalysis, and climate models (CMIP5 and CMIP6), plotted against (a–g) precipitation (P, mm month^{-1}); (h–n) surface shortwave radiation (RDN, W m^{-2}); and (o–u) leaf area index (LAI, m^2 m^{-2}). Satellite ET data are plotted against P from CHIRPS, RDN from CLARA-A1 and LAI from MODIS; ERA5 and climate model ET are plotted against ERA5 and model P, RDN and LAI, respectively. Data are from 2003 to 2013, with the exception of CMIP5, where the data are from 1994–2004. Note that the axes do not start at zero.

References

Gloor, M., Brienen, R. J. W., Galbraith, D., Feldpausch, T. R., Schöngart, J., Guyot, J. L., Espinoza, J. C., Lloyd, J. & Phillips, O. L. 2013. Intensification of the Amazon hydrological cycle over the last two decades. *Geophysical Research Letters,* 40**,** 1729-1733.