The authors present a comparison of drought metrics, calculated with different rainfall products. The study region is focused on the Amazon basin, and an extrapolation is made of aboveground forest carbon loss from drought. The authors end with a message that evaluation of drought through an ensemble is better. I think the comparison of rainfall products and evaluation of drought metrics could be useful, especially if it is more developed in the revision. This section could use some more analysis, especially with respect to defining anomalies per pixel location rather than absolute thresholds. However the section concerning the extrapolation of forest carbon loss from drought is a large overreach and does not help advance the state of the science. Please see the following general comments, and line comments.

- *We thank the reviewer for his/her very constructive feedback and detailed assessment of our study. We have addressed all comments below.*

General comments:

Carbon loss from drought - I will start with my strongest objection to this study, which is the extrapolation of forest carbon loss from drought. Accurate estimation of tropical forest carbon loss from drought is a highly sought after goal for tropical ecosystem ecology, but the methods this study uses are not robust or defensible in the present day. The standing biomass and forest sensitivity to drought differs dramatically across Amazonia. This point is even acknowledged (Line 435) in the manuscript. This study does not present any new field data to evaluate this very simplistic empirical relationship (from Lewis 2011), and therefore this study does not have the substance to make these claims. Even Lewis (2011) states this is a first approximation approach and does not include any goodness of fit statistics, the number of plots used to derive this estimate, or even specific information about which RAINFOR plots were included. Lewis extrapolated the relationship beyond the MCWD observed within the RAINFOR plot network from the 2005 drought through the 2010 drought to produce a quick estimate of carbon loss. In this study, the simplistic linear relationship is extrapolated even further beyond the original Lewis 2011 extrapolation. Even if this original relationship was remotely accurate for the 2005 drought, there is no evidence that it was accurate for subsequent droughts in 2010 (or 2015/16). It is difficult to make these forest carbon loss estimates regarding the 2015/16 drought without new field observations and validation, therefore I do not agree that the AGB loss estimates presented here are justifiable and object to their inclusion.

- *We appreciate the reviewers' very comprehensive comments. We agree that we overlooked the study of Feldpausch et a. 2016 which shows that the 2005 AGB-MCWD relationship cannot be applied for 2010 and no evidence exists which would justify the application of the relationship ship for 2015/2016. Hence, we will remove the according estimation of the drought impact on AGB for 2010 and 2015/16, but keep the estimate for 2005. For 2010 and 2015/2016 we will focus on the comparison of the drought indices instead of the AGB estimates.*
- *Generally, the point of our study is not to give better estimates of AGB loss during drought, but rather show how the choice and version of a climate (forcing) dataset also can have large influences on the drought impact and representation. In addition, we wanted to highlight that despite having better satellites and more sophisticated techniques when estimating gridded*

rainfall maps uncertainty rather increases than decreases. We will add a better explanation about this into the discussion.

Next, it is worth noting that a large-scale squall line also crossed the Amazon basin during the period of measurements presented in the original Phillips 2009 Science paper. This was estimated to have killed hundreds of millions of trees (Negrón-Juárez et al., 2010 Geophysical Research Letters), so even the empirical AGB~MCWD loss relationship presented in Lewis 2011 has a heavy bias from wind mortality. I strongly urge the authors to drop this aspect of the manuscript. Estimating Amazonian forest carbon loss from drought has long been a difficult endeavour, and many groups have been physically collecting field observations to quantify this. I worry this aspect of the study adds more noise than value to the current state of the science.

- *We thank the reviewer for this detailed comment. As stated in the comment above we agree with the reviewer and will remove the AGB-MCWD relationship for 2010.*

Defining drought - I think the evaluation of different precipitation datasets concerning the drought is mostly fine and could be useful. However the way drought is defined here is a bit simplistic, especially regarding the MCWD anomaly. The mean annual precip spans from 3500 mm + in the northwest Amazon to less than 1700 mm in the southeastern peripheries. I think it is difficult to justify a definition of drought based on absolute thresholds for the MCWD anomaly. The northwest Amazon rarely experiences a dry season, whereas the southeast Amazon does not receive rainfall for more than half the year. Forests are adapted to some level of water stress, which is why simple absolute thresholds are unlikely to characterize vegetation water stress. Assessing drought anomalies based on the number of standard deviations (calculated per pixellocation) is one commonly used way to assess drought with respect to the baseline climate and interannual variability of precipitation. Absolute thresholds (e.g. MCWD >25) vs. relative anomalies (e.g. MCWD > 2 standard deviations). The older papers using MCWD (e.g. Aragão et al., 2007) used a fixed value because there was not enough information at the time of actual ET. Now it is well understood that actual ET can vary substantially across the Amazon and has seasonality in most regions. It no longer makes sense to use a fixed value of ET for both the everwet northwest Amazon and the seasonally dry southwest Amazon. I suggest the authors could use newer spatially resolved ET estimates such as from GLEAM, MODIS MOD16, Fluxcom, etc.

- *While we generally agree with the reviewer that precipitation is very heterogeneously distributed across the Amazon rainforest, the constant ET of 100 is still being used in very recent publications (e.g. Flack-Prain et al. 2019, Biogeosciences; Koch et al. 2021 Earth's Future). As this study was not intended to develop a better drought index, but rather on the comparisons, we would keep the simple ET = 100mm.*
- *We already also did some analyses in the SI using ERA5 ET instead of constant ET=100 which did not affect affected areas too much. We will highlight these differences more in the main study and provide further figures in the SI.*
- *While initially, absolute thresholds are useful for deriving the absolute impact of AGB change, we agree switching to relative thresholds is more meaningful.*

The comparison of precipitation products and drought metrics could be a useful contribution, however this is currently muddled by putting all the estimates together in an ensemble. I suggest the authors focus on presenting a more organized comparison of (1) precipitation products, and (2) drought metrics. What is the justification for using an ensemble of precipitation datasets? Why is this better than using the best evaluated precipitation dataset? Consider the timing of the development of these products. Some of them have been operational for over 20 years. Statistical methods, data assimilation and climate reanalysis models have improved dramatically since then. I think it is difficult to argue that an ensemble method is better, especially when including where a coarse resolution earlier generation product (e.g. GPCC) has as much vote as the latest generation of products (e.g. ERA5, GPM IMERG6).

- *We do not fully understand the reviewers critique regarding our approach. We use the term "ensemble" to reflect a collection of datasets that have overlapping spatial and temporal resolutions. We still consider each dataset individually. As mentioned in the text, the scope of our study was to conduct "a systematic analysis of how the most frequently used precipitation datasets differ regarding the spatial extent, location and severity of recent extreme drought events". Obviously, we were not clear enough about this scope and will make this clearer throughout the text.*

Other comments

There are a number of typos in both the main text and figures. Some of these are highlighted in the line comments.
There are far too many acronyms in this manuscript. For example, is CHR really a useful shortening of the CHIRPS? Each new acronym makes the manuscript more difficult to read. I suggest limiting the usage of acronyms to the absolute minimum. Wherever possible, use established acronyms such as TRMM. Making up new acronyms of acronyms (TR6, TR7) is confusing and will not help readers comprehend the manuscript. A manuscript of this length does not need additional acronyms to make it shorter.

- *We fully agree and will use the official acronyms instead of making up new ones.*

Section comments:
L30: This should be MCWD > 25 mm, no? Also the climatological mean MCWD across Amazonia is quite large. I don't think it makes sense to use a single value to define drought (~25 mm). MCWD >= 25 mm in the southeast Amazon does not indicate drought.

- *Similarly to Lewis et al. 2011 we wanted to use the negative definition of DeltaMCWD, so in this case DeltaMCWD < -25mm would be correct.*
- *However, as already stated above we agree to switching to relative thresholds and will rephrase this part accordingly.*

L 170: The wet season starts at different times of the year across the Amazon. How is the choice of starting the hydrological year determined?

- *Similar to Phillips et al. 2009 and Lewis et al. 2011 we selected the 1st October as the onset of the hydrological year for each location in the Amazon. We will add this more clearly to the methods.*

L 173: I am not sure Delta MCWD is a good abbreviation for the anomaly of MCWD. This can easily be taken as just the change in MCWD between two time periods, but that's not exactly what the anomaly is during a drought. Perhaps it's better to spell it out as the " MCWD anomaly".

- *We thank the referee for this suggestion and will use the term MCWD anomaly throughout the text and figures.*

L 176: Removing the drought years causes bias. There are three droughts in the span of 15 years, so these are not rare events. Just because Lewis 2011 used a method, does not mean it is defensible in the present day.

- *see subsequent comment*

L 185: Climatologies are typically calculated from 30 year periods. Most of the data products have at least 20 years of duration, if not closer to 40. The selection of years to remove is subjective and removing the years with anomalously low rainfall will bias the standard deviation to be artificially small.

- *Regarding L176 and L185, we agree with the reviewer that removing the 3 extreme events may cause a bias. We will include the baseline years 2005, 2010 and 2016 in our MCWD calculation to avoid this bias.*

L208: Be consistent in treating MCWD as either a positive or negative quantity.

- *We will correct any inconsistent use of MCWD throughout the manuscript.*

L215+: I reject the underlying basis for the empirical carbon loss estimate from Lewis (2011).

- *We accept this rejection and do not give estimates of carbon losses for years other than 2005.*

L229: MCWD is misspelled L295: It is difficult for rainfall products to correctly estimate rainfall near the foothills of the Andes. Also some areas have very little ground information for each product's bias correction algorithm. It might be worth getting into this to describe more deeply why the products disagree, and where.

- *We thank the reviewer for mentioning this very important point. We will get into more detail at this point about why datasets disagree near the Andes.*

L333: I would note that many studies no longer use the fixed estimate of 100 mm. I believe some have used Stephenson (1998 Journal of Biogeography) as a reference for the development of the MCWD metric.

- *We will mention Stephenson (1998 Journal of Biogeography) as a reference for MCWD and we will also mention that there are better alternatives for ET instead of assuming constant ET =*

100mm/month. However, there are still quite some studies that are using the constant ET= 100mm/month approximation which we will also highlight at this point.

L357: Using a better estimate of "actual ET" might reflect the impact of VPD. I would say this is a limitation of using a fixed 100mm value for ET in the MCWD calculation.

- *This is a good point and we will include this in our discussion.*

L426: Indeed, this is another reason to drop the extrapolated carbon loss estimates.

- *See above sections.*

L453: I don't think the case for assessing drought with an ensemble is made clear.

- *We will rephrase this to make it more clear (see also comment below)*

Why is it not better to just use the product that has the lowest RMSE in the region of interest?

- *We again want to highlight that the purpose of this study is not to find the best dataset for locations at which we have exact measurements and can evaluate RMSE, but to give a broad picture of how different precipitation datasets represent drought stress across the complete basin. We will rephrase some parts of the manuscript to make this more clear.*

L458: The code in the repo looks to be incomplete. Ideally the complete code for analysis and figures should be hosted prior to the review process. An incomplete repository hinders the review process.

- *We will put all the files and scripts on the repository so that they can be easily reproduced.*

Figure 1: Is this MCWD, or anomalies of MCWD?
Figure 2: Why is WAT not included in panel C?
Figure 3: This is a useful figure. It might be useful to add another two columns indicating where the satellite based products agree, and where the climate reanalysis modeled products agree.
Figure 4: Is "PA" (y-axis label) supposed to be "RAI"?
Figure 5: Is "PA" (y-axis label) supposed to be "RAI"? The delta MCWD supposed to be the Anomaly of MCWD? Might be better to spell this out.
Figure 6: I suggest removing this aspect of the study, and this figure.
Table 1: I suggest dropping the abbreviations of abbreviations, and adding a column about how the product is derived (e.g. Remote sensing, interpolation of ground data, atmospheric process model, etc).
Table 2: RAI?

- *We thank the reviewer for the in-depth checking of our figures and will fix all the listed issues.*