Review #2:

In this analysis, the authors attempted to provide a new alternative to track the spatial-temporal variations in water availability in Rio Santa basin (RSB). They acknowledged the limitation of in situ and remote sensing-based rainfall datasets for complex terrain, i.e., the problematic quality and temporal consistence. While it is surprising that the authors found that satellite-derived vegetation greenness (also phenology) was coupled well with recent changes in rainfall (CHIRPS, a combination of satellite and rain gauge data). The authors proposed a "bucket" model to better fit vegetation phenology derived from MODIS NDVI, the concerns also raised in the criteria in extracting SOS/EOS. Some additional serious issues related to the methods and result interpretation, the organization of discussion would weaken the reliability and implication of this study. A major revision is therefore recommended.

We thank the reviewer for the evaluation of our manuscript and the insightful comments.

First, I have noticed that the authors claimed that both in situ and remote sensing precipitation datasets were questionable in signifying the changes in the timing and intensity of the wet season (e.g., lines: 4-7). However, it is quite strange that the authors used such dataset to demonstrate that the pattern of precipitation occurrence and the seasonality of vegetation indices are tightly coupled (e.g., lines: 10-11). Moreover, the authors used gridded precipitation data as a proxy of water availability (e.g., lines: 179-189). I would like the authors to rephrase these sentences and make the abstract more logical.

We agree with the reviewer that the strengths and limitations in using rainfall datasets for variability and trend analyses should be clarified and we made amendments to the manuscript to that effect. Specifically, we pointed out that in-situ rainfall data (i.e. weather station data) has limited spatial information in a complex terrain environment such as the Rio Santa basin. Merged satellite-based precipitation data on the other hand can suffer from spatio-temporal inconsistencies related to sensor overpass times, sensor changes, as well as biases introduced by the retrieval algorithm or lack in reference in-situ data. Cross-validation of trends with independent datasets hence considerably increases trend robustness. Other precipitation metrics that are less sensitive to temporal inconsistencies are likely to be less affected, among these are co-variabilities between precipitation and vegetation as indicated in Fig.2a or rainy season onset/retreat metrics. To clarify this, we added a few sentences to the Introduction section stating why we did not primarily exploit rainfall datasets. The use of CHIRPS here is justified by the fact that CHIRPS data has been used for rainfall trend studies in other regions of the Andes before (e.g. Segura et al., 2019, Torres-Batlló & Martí-Cardona, 2020), and that it is temporally more consistent than IMERG data, where the weighting between microwave and infrared data dynamically changes over time and therefore can lead to strong (and potentially unrealistic) trends as seen in Fig.3 IMERG trend. In the revised manuscript, we clearly stated that CHIRPS data for the RSB is in fact able to capture seasonal variability in rainfall (as shown by the bucket model, i.e. Fig.5) and to some extent represents trends in rainfall coinciding with sub-seasonal trends in vegetation greenness in the observed period (as shown in the small inset plots in Fig.4i and l).

Second, some key messages are missing in the method section.

i) How did the authors reconcile the NDVI, precipitation and soil moisture data with different spatial resolution? It should be noted that the topographical issues in the studied mountain areas should not be ignored;

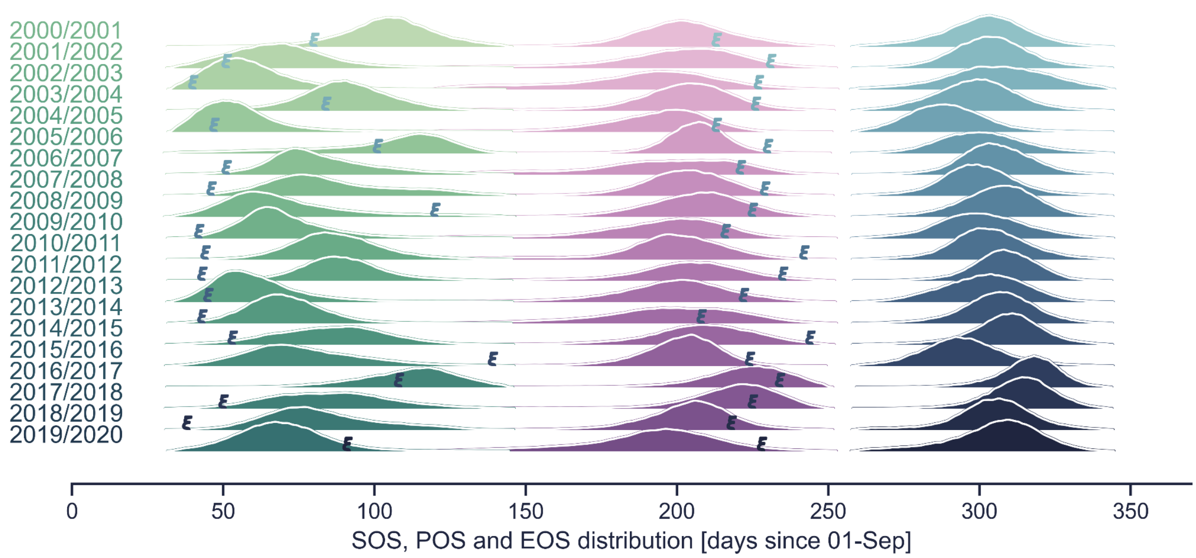
In fact, all analyses are based on either native resolution of the datasets or spatial averaging over the entire basin as stated in the dataset description section. The only analysis where rainfall and vegetation indices data were jointly used for sub-domain calculations is the lagged correlation, where we state that "… we compared each VI pixel with the CHIRPS pixel intersecting it by using a nearest neighbor approach." (lines 157-158). Consequently, no interpolation was applied to either of the datasets and no additional errors or uncertainties related to the topography in the region should be introduced by our analysis. We added a sentence to the associated methods section to clarify and avoid the reader misinterpreting.

ii) The authors should provide more details related to the lagged correlation, e.g., the mathematic implementation of a cross correlation function;

We added an explanation of the lagged correlation, along with the corresponding equations.

iii) It is interesting that why evapotranspiration data in the "bucket" model (e.g., Eq. 1) is set as constant over the study period? In other words, the seasonal variations in soil moisture are solely determined by precipitation? Then, where is the advantage of "bucket" model against the seasonal rainfall data (e.g., Liebmann and Marengo, 2001, also Figure A1)? A better criterion to extract SOS/EOS?

Evapotranspiration is set to a constant value for the sake of simplicity. The aim of the bucket model is to crudely represent water retention in the soil in the most basic way, which however allows to estimate plant available water rather than rainfall accumulation alone as with the methods presented in Appendix A. This allows to represent the observed asymmetry in vegetation lag at the start and the end of the rainy season, i.e. a small lag in greening after rainfall onset but a larger lag for vegetation EOS at the end of the rainy season, when rainfall ceases but soil water is still abundant. For additional explanation we added a modified version of Fig. 5 where the rainy season onset and retreat derived by the method of Liebmann & Marengo (2001) instead of the bucket model is presented. For simplicity, the KDE's represent the whole Rio Santa basin (not distinguishing between Coordillera Negra & Blanca as in the original Fig.5).



iv) Similarly, the authors should state the rationale of applying two specific thresholds to define the simulated SOS/EOS (e.g., 0.2 and 0.35 m3/m3, lines: 169-172). Are these thresholds specifically optimized for the NDVI-based SOS/EOS?

Yes, the two thresholds were determined in order to highlight the congruence between the onset and retreat of the wet season (as inferred from precipitation data and evapotranspiration estimates) and the onset/end of the vegetation period as inferred from NDVI. We acknowledge that there is a certain arbitrariness in this choice, as the two thresholds (along with the two parameters needed to define the minimum and maximum value of the soil water bucket, viz. 0.05 and 0.5 m3/m) are plausible but not measured physical properties of the soils and vegetation types occurring in the region. We would like to reiterate here that the purpose of this exercise was not to provide a realistic simulation of the soil moisture balance and its seasonal variations at the pixel level, but rather to show i) that accounting for a lag in the response of the soil-vegetation system to the onset and retreat of the rainy season can explain the seasonal patterns derived from the NDVI data aggregated at regional level, and ii) that for this purpose CHIRPS provides a reasonable representation of the precipitation regime. Similar considerations motivate our choice of setting evapotranspiration constant to 2 mm/d in this context.

We are aware that the agreement between simulated and NDVI-derived metrics of land surface phenology could be improved by using a more realistic formulation and parameterization of the soil moisture balance. To make this clearer for the reader, we introduced amendments to the corresponding sections (Introduction and section 2.6), stating the rationale for introducing the bucket model.

Third, the authors realized the NDVI signals were lagged behind the precipitation (e.g., Figure 2.a). i) Why not presented the variations of precipitation and NDVI after few months (instead of Figure 3). In theory, it could be able to support the coherence of SOS/EOS inferred from vegetation index and precipitation data. Unfortunately, it looks like the NDVI is always greening even take the lagged months into consideration.

We apologise if we do not correctly understand the point of improvement the reviewer is suggesting here. The coherence of NDVI/precip SOS/EOS is shown (based on entirely independent models) and discussed in Figure 5 and could not be visualised based on annual trend plots as shown in Figure 3, even if a 1 or 2 month lag was introduced to the NDVI. As described in the manuscript, the purpose of Figure 3 is exclusively to present average regional trends, and the difficulty in reconciling datasets based on this metric. We then go on to illustrate that there is however good coherence in SOS/EOS metrics between NDVI/CHIRPS and some correspondence specifically in the late rainy season trend, allowing us to identify this as a robust feature of sub-seasonal increased water availability with direct effect on plant greening.

ii) As is shown in Figure 6, the lags between SOS derived from MODIS NDVI and CHIRPS rainfall data and that for EOS were the same?

As the reviewer mentioned before, section 2.5 describing the lag correlation was incomplete and was updated in the new manuscript version. We agree that interpretation of the maps shown in Fig.6d) and h) may have been difficult. For explanation, the “best lag” values shown in Fig 6d) show the median of 20 years of optimized lags. These were calculated as follows: for each gridpoint, the seasonal evolution of the NDVI and precipitation were selected similar as done in Fig. 2 for all seasons for the spatially aggregated values; next, we determined by how many days the seasonal curve of one variable had to be shifted to obtain the highest Pearson's r value. The main outcome of this analysis is a qualitative representation of the effect of topographic/spatial features on the rainfall-vegetation relations (as described in lines 217-224).