In the following we use R3C1 (etc) to refer to comment 1 (C1) by referee 3 (R3).

Anonymous Referee #3

R3C1: This study tries to partition the inter-annual variability in precipitation (P), i.e., the source term in terrestrial water cycle, into variabilities in three sink terms in terrestrial water cycle (ET, Q, ΔS), and then to relate the partitioning of variabilities to various factors like temperature, aridity, and storage capacity. I think this type of study at global scale is rather new, if not first of its kind at global scale, and thus very interesting to the hydrology community. This is the case mostly because there has been a lack of "hydrologic reanalysis" (CDR) for such kind of analysis in the first place. At the same time, this effort couldn't fully answer many of the questions set forth at the beginning, leaving perhaps "more questions than answers" (as phrased by another referee). The authors have done a solid amount of thorough analysis and experiments toward the questions of interest and these analyses are also well designed too.

Overall I consider this manuscript of good quality, both scientifically and technically, and thus publishable in HESS with several concerns addressed.

Response: We agree that this is a first-of-its-kind study and thank the referee for the encouraging positive comments on the manuscript.

R3C2: My primary concern is there is a lack of general "signal-to-noise" discussions to better inform readers to what extent the findings are significant signals from the underlying data (CDR, Zhang et al., 2018) and how much of it could be due to data uncertainties (or possible artifacts due to how the data is produced). For example, the ET products that went into the CDR (satellite products, reanalysis, etc.) share some similarity in their production methods (e.g., Penman-Monteith or Priestley-Taylor type of schemes). Such similarity may limit the variability of ET in CDR. Of course, the plants do apply a strong filter on the inter-annual variability based on their survival need. Such uncertainty analysis may be difficult but I think some qualitative and general assessment would be very beneficial.

Response: The CDR uses a formal data assimilation scheme based on mass balance that weights the various inputs, and thereby produces uncertainty estimates for each variable (P, E, Q, ΔS). The original paper (Zhang et al., 2018 HESS) includes a formal assessment of the sensitivity of P, E, Q over large regions (continents, basins) using the coefficient of variation (see original Figures 2, 3, 4, 5, 6, 7 in Zheng et al., 2018 HESS). We actually followed from that work and used those uncertainty estimates (lines 122-130) to identify and mask out regions where we judged the uncertainty to be large relative to the magnitude of the fluxes. This screening procedure removed most grid-boxes from the Himalayas, Sahara Desert and Greenland (see Fig. 1).

Secondly, while it is true that some of the products might share similarity in producing, for example, *E* (Penman-Monteith, Priestley-Taylor as the examples noted by the reviewer) the data assimilation is a comprehensive approach that includes all available estimates of *P*, *E*, *Q* and Δ*S* at each grid box. With mass balance enforced, the CDR estimates represent a composite product that is designed to avoid bias of the type described by the reviewer as much as possible by using all available estimates of the hydrologic fluxes. As we have described in a response to Reviewer 2 (see R2C3), the CDR has been extensively validated in the original publication. In that context, our goal was not to assess the CDR, but rather to use it for this "first-of-a-kind" study on the sources and sinks of inter-annual hydrologic variability.

In summary, with the many individual validations of the CDR in the original paper (Zhang et al., 2018 HESS) augmented by those in our manuscript, our results are based on the best available hydrologic reanalyses. In terms of the remaining uncertainty from the CDR data, this is beyond the scope of the current study. Despite that, the general approach in our manuscript will remain and the results will be fine-tuned over the coming years as the hydrologic community develops and uses their own reanalyses. We will add words to that effect in the revised version of the manuscript.

R3C3: Also, at the scale of the CDR (0.5 degree), I would say the partitioning is more complicated than just a result of several factors. The horizontal transport of water, seasonality, local water use, etc., can add a lot of noise. I wouldn't say it is not possible to do it at 0.5 degree, but it would probably be less noisy at a slightly coarser scale. Also, there could be much more controlling factors for the partitioning than being investigated, e.g., land cover/land use, LAI, topography.

Response: We agree with the reviewer that the partitioning is complex and could be related to the other factors, e.g., land cover/land use, LAI and horizontal transport of water due to topography, etc. In this first-of-a-kind analysis we chose to focus on the zero'th order physical factors (storage capacity, snow/ice) at the CDR data resolution (0.5 degree), but we fully expect more detailed analysis to follow, e.g., vegetation plant-based variables as discussed by the reviewer.

R3C4: Finally, given that this study does tend to raise more questions than answers, I feel the authors should provide some more insights on what we can do from the analysis and findings in this study. What can we do with the numbers concluded here? Validating models? Improving single models like Budyko? Hydrologic/water risk analysis? Climate system behavior/sensitivity and hydrologic impacts of climate changes? And how can we improve our understanding in the future? What kind of new data at what scales would be critical to answering such questions? I feel this paper is incomplete without offering some of such insights.

Response: We thank the reviewer for the constructive suggestion on the insights of this study. This is awkward – what the reviewer is asking for is an extended discussion while reviewer 2 has asked for less discussion. In the revision we will try our best to find a balance and set out what can be learnt. To respond in more detail, what we have learnt from undertaking the study

is that; (i) partitioning of hydrologic variability does not follow partitioning of the mean, (ii) the long-ignored covariances play a critical role in hydrologic partitioning, especially in biologically productive environments (aridity index ~1), (iii) and because of those covariances there will be no simple translation of changes in the variability of P into changes in the variability of E, Q, ΔS. We also expect that in future we will be able to extract generic signatures of hydrologic variability (e.g., Fig. 8) that can be used to assess the simulation of variability in models. In response to this important point made by the reviewer we intend to carefully revise the discussion and add some of the potential implications as requested by the reviewer.