

**In the following we use R2C1 (etc) to refer to comment 1 (C1) by referee 2 (R2).**

**Dr René Orth**

R2C1: Review of Dongqin and Roderick “Inter-annual variability of the global terrestrial cycle”  
This study investigates the propagation of precipitation variability into the water cycle, i.e. into variations of runoff, evapotranspiration, and of storage changes. The authors show that this is mostly controlled by temperature (in wet regions), long-term aridity (in transitional regions), and by soil water storage capacity (in dry regions). Further, the results illustrate that the corresponding partitioning is different from the partitioning of mean precipitation into the means of these water cycle variables.

---

Recommendation: I think the paper requires major revisions.

The analysis is very interesting and provides new and fundamental insights into large-scale land surface hydrology. Related variability analyses are still not commonly done due to a lack of reliable data and underlying theory. This study can foster theory development in this area, and it underlines the importance of continuous improvement of the just-emerging global hydrological re-analysis datasets. Therefore I would be happy to see it published in HESS, but after some general revisions.

Response: We thank Dr René Orth for the evaluation and comments on the manuscript.

R2C2: (1) Next to the consideration of the soil water storage capacity and the mean temperature to explain variations in the partitioning of precipitation variability, I am missing the inclusion of vegetation type as an explanatory variable. It might have strong implications on evapotranspiration variability, and therefore also on runoff and storage variabilities.

Response: We agree with Dr René Orth that the inter-annual variability might be related to the other factors, e.g., vegetation type. However, given the fact that this is a new approach and the research is exploratory, we focused on relating the inter-annual variability with the most general hydrologic factors (i.e., the air temperature as a surrogate for snow/ice and water storage capacity). We expect to extend the current work to a more complete analysis (e.g., relation to vegetation) in future research.

R2C3: (2) I agree with the authors that comprehensive hydrological reanalysis datasets are lacking, and the CDR dataset is an important contribution in that respect. Further, I appreciate the effort they make to validate the applicability of the dataset in the context of this study. However, also the CDR dataset is (necessarily) based on a model and hence it is not clear that the reported relationships are operating in nature, and not only in this model. To address this

issue, I would like to see the key analyses from this study repeated with the state-of-the-art ERA5 reanalysis, which should be superior to ERA-Interim also in terms of land surface representation.

Response: One of the reasons for developing hydrologic reanalyses like the CDR (and hopefully other forthcoming hydrologic reanalyses) is that the broader hydrologic community were not satisfied with atmospheric reanalyses (e.g., ERA/NCEP). In that context, CDR is “hydrologic-centric” as opposed to ERA which is “atmospheric-centric”. In creating the original CDR, they actually used evaporation from the ERA-interim (as one of 8 different  $E$  products, see Table 1 in Zhang et al. 2018, HESS). To replace  $E$  from ERA-interim with that from ERA5 would require us to completely re-do the CDR data assimilation but that is well beyond the scope of this work.

In terms of relating the CDR to the real world there is ample evidence that it is suitable for the analysis conducted here including:

- (i) The enforcement of basic hydrologic concepts (mass balance).
- (ii) The numerous tests of CDR reported in the original Zhang et al 2018 HESS publication (that are summarized on lines 134-139 of the HESSD manuscript). Those tests include a (successful) comparison of CDR runoff to observations of monthly runoff at 165 medium size basins and 862 small basins. In fact, the assessment of CDR in the original paper was quite comprehensive as you would expect.
- (iii) We have augmented those extensive original tests by independently comparing monthly  $E$  with FLUXNET tower data at 32 sites which confirmed that the CDR captured the general seasonal cycle in both  $P$  and  $E$  at those 32 sites (Fig. S2, S3, S4, Table S1). We also used the same FLUXNET data to compare the variability in  $P$  with variability in  $E$  (Fig. S5).
- (iv) We further compared CDR  $E$  with two gridded  $E$  databases that are not included in the source databases of CDR (LandFluxEval, MPI, see lines 157-163 in the manuscript and Fig. S6, S7) and the comparison was satisfactory.
- (v) We compared how the standard deviation for  $E$  and the mean for  $E$  are related in the CDR (Fig. 6) and compared that with the same relations in LandFluxEval and MPI (Fig. S8). Those two comparisons were satisfactory.
- (vi) The mean water cycle ( $P$ ,  $E$ ,  $Q$ ) in CDR was shown to be consistent with the long-standing Budyko framework (Fig. 3).
- (vii) The CDR data were consistent with the Koster & Suarez (1999) theory in the limit of sites that have limited water storage (Fig. 7).

That is a very comprehensive assessment.

In summary, there will always be new databases for comparison but the above tests led us to conclude that the CDR was at least suitable for an exploratory analysis of water cycle variances using the completely new “variance balance approach” (Eqn 2) described in the manuscript.

R2C4: (3) I appreciate the idea of investigating the influence of the soil water storage capacity and the mean temperature on the variability partitioning. However, I think parts of the conclusions drawn by the authors from Figures 8-10 are not supported by the data. For example, I cannot see in Figure 10 that the temperature influence is particularly strong in very wet regions. Rather, to me it seems to be strong in moderately wet and dry regions (Fig 10b,d,f,h,j,l,n,p). Further, also the aridity limit of 6 which the authors suggest in their interpretation of the results in Figure 9, is arbitrary and not supported by the actual results. Storage capacity is obviously having an influence already for aridity values above 2-3 (Fig. 9b,c,f,j,k). Overall, in these Figures there are many interesting patterns but the authors focus only on few sub-plots and limit their interpretation to these. Therefore, I suggest to either show less information/sub-plots there, or to develop explanations also for patterns emerging within other sub-plots.

Response: We accept that Fig. 10 is hard to interpret. On reading the reviewers comments and going over the manuscript we realize the problem was that we did not explicitly indicate the relevant panels (i.e., a, b, c, ....). That was an oversight correctly identified by the reviewer. In general, the data in Fig. 10 was not particularly revealing (i.e., a negative result) but we actually focused the discussion to the first two columns but we did not identify them properly. In response, we propose to replace the original text with the following:

*“To understand the potential role of snow/ice in modifying the variance partitioning, we repeat the previous analysis (Fig. 9) but here we use the mean annual air temperature ( $\overline{T}_a$ ) to colour the grid-cells to crudely identify the presence of snow/ice (Fig. 10). Most of the variations at polar latitudes in the northern hemisphere (panels in the first and second columns of Fig. 10) is associated with low air temperature (e.g.,  $\overline{T}_a < 0$  °C), making the results associated with high air temperature (e.g.,  $\overline{T}_a > 10$  °C in third and fourth columns of Fig. 10) show less scatters. That pattern is particularly obvious in extremely wet environment ( $\overline{E}_o/\overline{P} \leq 0.5$ ), where the ratio  $\sigma_Q^2/\sigma_P^2$  is close to 1.0 when  $\overline{T}_a$  is high (e.g.,  $\overline{T}_a > 10$  °C, Fig. 10gh) but shows lots of scatters when  $\overline{T}_a$  is low (e.g.,  $\overline{T}_a < 0$  °C, Fig. 10ef). This indicates that in extremely wet environment, when  $\overline{T}_a$  is high,  $\sigma_P^2$  is almost completely partitioned into  $\sigma_Q^2$  (e.g.,  $\overline{E}_o/\overline{P} \leq 0.5$  and  $\overline{T}_a > 10$  °C in the third and fourth columns of Fig. 10). However, when  $\overline{T}_a$  is low in extremely wet environment, there are substantial variations in all variance-covariance components (e.g.,  $\overline{E}_o/\overline{P} \leq 0.5$  and  $\overline{T}_a < 0$  °C in the first and second columns of Fig. 10). That result indicates the complexity of variance partitioning associated with the presence of snow/ice.”*

R2C5: (4) The paper contains (too) many figures, which is diluting the main message(s), I feel. For example, Figures 1 and 2 could be merged, Figure 5 could be moved to the supplementary material, Figure 13 could be merged into Figure 8. The authors might have further ideas to reduce the amount of figures. Moreover, I do not really understand the difference between Figures 7 and 8, and why both are needed.

I do not wish to remain anonymous - René Orth.

Response: We respect the reviewer's opinion that we have too many figures – this is always a hard balance to get right to everyone's satisfaction. We can easily combine Fig. 1 and 2 as suggested, or alternatively, we could move Fig. 1 to the supporting material. We can also move Fig. 5 to the supporting material as suggested. However, we do not think Fig. 13 should be merged into Fig. 8 since the two figures belong to different sections (Fig. 8 for the relation between variance partitioning and aridity section, Fig. 13 for the case study section). Fig. 7 is a direct link to previous work while Fig. 8 is the variance partitioning in the CDR database. Hence while these two figures are similar, they make separate independent contributions to the manuscript.

---

Specific comments:

R2C6: line 8: Equation 2 not introduced yet line 13: It should be 'variabilities'.

Response: We will modify these texts in the revised version of manuscript. Thanks.

R2C7: line 15: Some word is missing towards the end of the line

Response: We have checked line 15 and did not find missing words?

R2C8: lines 35-39: Orth & Destouni (2018) might be relevant in this context and could be cited.

Response: The reference will be cited in the revised manuscript.

R2C9: line 37: Not sure I get the point here.

Response: We mean that droughts and floods are typical extremes but that hydrologic variability encompasses more than just droughts and floods, i.e., hydrologic variability occurs across all time-space scales.

R2C10: lines 106-118: Please clarify that what you are determining here is actually not the soil water storage capacity, but rather the active range within which the soil moisture varies.

Response: Yes, exactly. We can modify the text to make this explicit.

R2C11: lines 157-163: I would recommend to replace the LandFluxEVAL and the Jung et al. datasets with more recent gridded ET datasets such as the Jung et al. 2019 dataset and the GLEAM dataset (Martens et al. 2017).

Response: The reason we chose the LandFluxEVAL and MPI databases is that they are among the most widely used and validated *E* data that were also not used to develop the CDR database. We do not think adding a comparison to the latest GLEAM database would be as useful since an earlier version of GLEAM (v2a) was actually an input to the data assimilation scheme used to construct the CDR (see Table 1 in Zhang et al., 2018, HESS). Instead it would be better to re-do the CDR data assimilation incorporating the latest GLEAM database but that is well beyond the scope of this work. (Also see R2C3 for similar comments about ERA.) We could replace the MPI we used with the updated database (Jung et al., 2019) but we do not see how that would alter the results.

R2C12: line 180: Gudmundsson et al. (2016) might be relevant in this context and could be cited.

Response: The reference will be cited in the revised manuscript. Thanks.

R2C13: line 181: What is meant by seasonality here? I thought you are considering annual data? In general, I think the considered temporal and spatial scales and resolution need to be more clearly stated and motivated at the beginning of the manuscript. Also, the role of these decisions on the results could be discussed.

Response: Yes, we are using annual data. But we know that differences of the intra-year seasonal timing (phase) of precipitation and  $E_o$  do have an effect on the annual water balance (as per the seminal work by Chris Milly in the early 1990s.). We will make this more clear in the revised version.

R2C14: line 252/253: I could not find this discussion in section 5!? Would be important to explain these discrepancies, though.

Response: Thanks for pointing this oversight out. The underlying scientific issue here is that the original Koster and Suarez (1999) work assumed no change in long-term storage. In that sense the original results of Koster and Suarez (1999) can be seen as an upper limit and any variance in storage can only reduce the partitioning of variability in  $P$  to variability in  $E$  under dry conditions (Fig. 7). We will add a short discussion on this in the revised manuscript.

R2C15: line 327 & 333: ‘leaving very limited variance’ - not really true given your statement in lines 385-387

Response:

324 We show the  $P$ ,  $E$ ,  $Q$  and  $\Delta S$  time series along with the relevant variances and covariances in Fig. 12. Starting  
325 with the two dry sites, at the site with low storage capacity (Site 1), the time series shows that  $E$  closely follows  
326  $P$  leaving annual  $Q$  and  $\Delta S$  close to zero (Fig. 12a). The variance of  $P$  ( $\sigma_P^2 = 206.9 \text{ mm}^2$ ) is small and almost  
327 completely partitioned into the variance of  $E$  ( $\sigma_E^2 = 196.9 \text{ mm}^2$ ), leaving very limited variance for  $Q$ ,  $\Delta S$  and all  
328 three covariance components (Fig. 12b). At the site with high storage capacity (Site 2),  $E$ ,  $Q$  and  $\Delta S$  do not simply  
329 follow  $P$  (Fig. 12c). As a consequence, the variance of  $P$  ( $\sigma_P^2 = 2798.0 \text{ mm}^2$ ) is shared between  $E$  ( $\sigma_E^2 = 1150.2$   
330  $\text{mm}^2$ ),  $\Delta S$  ( $\sigma_{\Delta S}^2 = 800.5 \text{ mm}^2$ ) and their covariance component ( $2cov(E, \Delta S) = 538.4 \text{ mm}^2$ , Fig. 12d). Switching  
331 now to the remaining wet and hot site (Site 3),  $Q$  closely follows  $P$ , with  $\Delta S$  close to zero and  $E$  showing little  
332 inter-annual variation (Fig. 12e). The variance of  $P$  ( $\sigma_P^2 = 57374.4 \text{ mm}^2$ ) is relatively large and almost completely  
333 partitioned into the variance of  $Q$  ( $\sigma_Q^2 = 57296.4 \text{ mm}^2$ ), leaving very limited variance for  $E$  and  $\Delta S$  and the three  
334 covariance components (Fig. 12f). We also examined numerous other sites with similar extreme conditions as the  
335 three case study sites and found the same basic patterns as reported above.

The text here refers to the site-based case studies (line 327 – Fig. 12a – Site 1; line 333 – Fig. 12 f – Site 3) while the later text (lines 385-387) refers to the general pattern across all grid-boxes, i.e., Fig. 4. Perhaps we can correct this misunderstanding by rewriting lines 385-387 to indicate the relevant figures as follows:

“Hence we were initially surprised that the inter-annual variability of water storage change ( $\sigma_{\Delta S}^2$ ) is typically larger than the inter-annual variability of evapotranspiration ( $\sigma_E^2$ ) (cf. Fig. 4b and 4d). Moreover, the covariance components are also prominent and can be negative (Fig. 4efg), which means that it is possible for the variability in the sinks (e.g.,  $\sigma_Q^2$ ,  $\sigma_{\Delta S}^2$ ) can actually exceed the variability in the source ( $\sigma_P^2$ ) (Eq. 2).”

R2C16: lines 403-405: I cannot see this from Figure 8.

Response: Agreed. That was our mistake. The reference to Fig. 8 should be to Fig. 4 (global pattern of water cycle variability) and we will revise that in the revision.

R2C17: Section 5: Overall a bit lengthy with too much summarizing, I think. Could be shorter, and more concise.

Response: We will read and revise the Section 5 carefully and make it more concise accordingly in the revision as per the comments of both R2 and R3.

R2C18: Figure 3: Why are there data points outside the physically plausible range?

Response: We assume you mean points with  $E$  exceeding  $P$ ? This is possible in for example, regions with run-on, or irrigation. We have further investigated those points and also find that some of them come from the parts of Greenland that had not been masked out (Fig. 1). Non-steady state conditions, e.g. long-term changes in storage can also lead to  $E$  exceeding  $P$ .

R2C19: Figure 4: Many values seem to be cut at 10 as this is the end of the color bar. You could use log scale here for the color bar.

Response: Yes, the scale for  $P$  in Fig. 4a is saturated with the maximum value of the color bar 10000. The reason we chose 10000 as the limit was to show the patterns for both the relative high (e.g.,  $\sigma_P^2$ ,  $\sigma_Q^2$  and  $\sigma_{\Delta S}^2$ ) and low variabilities (e.g.,  $\sigma_E^2$ ,  $2\text{cov}(E, \Delta S)$ ). We can modify by using a log scale to address this comment.

R2C20: References:

Gudmundsson, L., P. Greve, and S. I. Seneviratne, 2016: The sensitivity of water availability to changes in the aridity index and other factors—A probabilistic analysis in the Budyko space, *Geophys. Res. Lett.* 43 (13), 6985-6994.

Jung, M., S. Koirala, U. Weber, K. Ichii, F. Gans, G. Camps-Valls, D. Papale, C. Schwalm, G. Tramontana, and M. Reichstein, 2019: The FLUXCOM ensemble of global land-atmosphere energy fluxes. *Scientific Data*, 6 (74).

Martens, B., D. G. Miralles, H. Lievens, R. van der Schalie, R. A. M. de Jeu, D. Fernández-Prieto, H. E. Beck, W. A. Dorigo, and N. E. C. Verhoest, 2017: GLEAM v3: satellite-based land evaporation and root-zone soil moisture, *Geosci. Model Dev.* 10, 1903–1925.

Orth, R., and G. Destouni, 2018: Drought reduces blue-water fluxes more strongly than green-water fluxes in Europe. *Nature Communications*, 9, 3602, doi: 10.1038/s41467-018-06013-7

Response: We appreciate Dr René Orth for listing all the reference mentioned above in the comments, and we will read and cite these reference accordingly in the revised manuscript. Thanks.