EGUsphere, referee comment RC2
https://doi.org/10.5194/egusphere-2022-769-RC2, 2022

Comment on egusphere-2022-769

Anonymous Referee #2

Referee comment on "Daytime-only-mean data enhances understanding of land-atmosphere coupling" by Zun Yin et al., EGUsphere, https://doi.org/10.5194/egusphere-2022-769-RC2, 2022

This paper demonstrates that the averaging approach employed by models (in this case, ERA5) in generating output diagnostics has an impact on what can be inferred from those diagnostics in the context of land-atmosphere coupling strength. Intuitively, such an impact makes perfect sense. I agree with the authors' call to modeling centers to provide relevant land-atmosphere coupling diagnostics at higher time resolution for improved analysis of land-atmosphere coupling.

All this being said, I must recommend major revision for this paper. The analysis strategy used is far from intuitive, and after reading the paper several times, I'm left unconvinced that the particular strategy used here is optimal (though I don't pretend to know what the optimal strategy is). It almost goes without saying that daytime-only data can get at the two-legged metric better than full-day or full-month data; still, I can't wrap my head around the idea that the quantile approach is the best way to tell us what we want to know (see comment 2 below).

1. Because the quantile analysis approach is not intuitive, further exposition in the Methods section would go a long way toward making this study more comprehensible. Perhaps the authors have spent so much time thinking about the analysis approach that it comes as second nature to them, but they should know that this won't be the case for the average reader. Significant additional explanation is needed. For example, I'm guessing that quantiles are based on all land (non land-ice) points across the globe. True? Please clarify. Also, are the quantiles computed separately for each season? If so, why are southern hemisphere JJA points mixed in with northern hemisphere JJA points in determining the quantiles? One would think that seasonal variations in the diagnostics would be hemisphere-specific.

2. Assuming that I do know what the authors are doing, I have some misgivings about what the quantile approach can tell us. Would consideration of only northern hemisphere extratropical points (probably a much cleaner approach, given seasonality) give the same results? Would a continental-scale analysis (e.g., North America only) give the same results? There's no way of knowing a priori; one can only speculate. Also, consider two highly hypothetical scenarios:

a) The TLM values produced with all three averaging approaches are perfectly valid (i.e., are perfectly consistent with each other) except over 20% of the Earth (defined by vegetation type, location on the globe, or whatever). In that 20%, the monthly averaging approach inappropriately assigns a very high coupling strength when the actual coupling strength is very low. In this hypothetical example, the monthly averaging approach would look very bad at the high extreme, as it should, but it would also look bad (20% off) everywhere else, when this example's assumptions say that it actually works just fine. This seems to be a basic limitation of the quantile approach.

b) In a separate hypothetical example, suppose that 80% of the globe experiences no land-atmosphere feedback of any relevance at all. In this case, quantile differences found between the averaging approaches within this lower 80% would have no practical meaning, and there'd be no point, e.g., in plotting quantile changes.

I'm not saying that these scenarios are realistic; I'm just saying that it's easy to come up with scenarios that call into question the understanding that can be gained from a quantile-based analysis. The authors should provide significant discussion about the limitations of dealing with quantiles like this.

3. I disagree with the conclusion on lines 234-236, in reference to the Koster et al. study. That study did not use the two-legged approach to quantify coupling; it simply quantified the impact of soil moisture variations on precipitation variability at the multi-day time-scale. For the particular coupling characterization it was after, the calculation was exact and was not limited in any way by daytime-only vs. all-day vs. multi-day considerations. The results of the present study are best considered in relation to studies that use the two-legged metric .

4. Section 2.4 came off as opaque to me. What does the "top 25% quantile" refer to – if it refers to the ACF values, why are the lower values being ignored? Why is the ratio of the sigmas relevant? What is meant by "numerator of the rho term"? Why is the relevance of the ratio of the N terms? Also, though I can kind of guess what are the authors getting at when they talk about signal attenuation in the first place, I can't be sure. A major rewrite is needed here.

5. The correlations in Figure 3 are undoubtedly statistically significant, but they are far from "high" (line 185) or even "moderately large" (line 191). Those in panels (a) and (b) indicate only a 10% explanation of variance, and those in the remaining panels indicate well less than half the variance explained. The text, though, presents these fields as clear indications that the authors have identified the main controls on various quantities ("Significant correlation coefficients suggest that our indicator adequately explains the attenuation…"). To be honest, I got very little out of Figure 3 and the associated discussion.

6. Would it be appropriate to at least mention that the daytime-only diagnostics may produce different results from midday-only diagnostics (e.g., 10AM-2PM)? Presumably not much coupling occurs at dusk and dawn. I'm not suggesting that midday diagnostics be examined in this paper; it's just that the overall problem of optimal averaging time goes beyond simply comparing all-day diagnostics to daytime-only diagnostics.