Review #1 of 'Indices of the Hadley circulation strength and associated circulation trends' by Pikovnik et al.

General comments:

This manuscript compares 8 measures of the HC strength derived from ERA5 and ERA-interim reanalysis datasets. Their main findings are that measures based on a single vertical level are more subject to uncertainty and inhomogeneity while measures based on spatial average or integration are more robust. They concluded that the measure of the average HC strength is best suited for studying variability and trends.

The comparison is interesting and the conclusions are pertinent. However, 7 out of the 8 measures are derived directly or indirectly from the zonal mean stream function, which explains the high correlations between measures. The one measure not derived from the stream function is deemed inadequate for this study and needs further refinement. Perhaps it would have been important to compare independent measures of the HC strength and quantify their relative relevance rather than the 7 measures proposed here as it is intuitive that capturing the HC by taking into account both its meridional and vertical extent would be more robust than from a single location.

We thank the Reviewer for their constructive comments. We hope that all their comments and concerns are adequately addressed in our responses and that the text modifications in the revised manuscript are appropriate. Figures 2-5 were made more concise by showing stream-function-based metrics 100 hPa apart. Furthermore, to reduce redundancy, Figures 2 and 4 now show trends of seasonal-mean HC instead of monthly-mean HC, following the argument presented in Waugh et al. (2018). In order to follow the terminology used by other authors (e.g. Solomon et. al, 2016; David and Briner, 2018; Waugh et al., 2018), we have changed “indices” and “measures” into “metrics” throughout the text, as well as in the title.

As the Reviewer points out, our suggestion that the metrics of the HC strength that take into account both the meridional and vertical extent of the global HC are overall better indices than the HC metrics based on local values may be intuitive. The time series of the stream function-based indices are aligned (Fig. 3 in the revised paper) and highly correlated (Fig. 5), however, the differences become important when one computes the HC trends and quantifies their uncertainties. This is the first lesson from our comparison of independent metrics of the HC strength. We applied the metrics from previous studies to bring our results into the context of the reported trends in the HC strength. We agree that the trends based on other independent metrics could be explored, such as the water vapor flow of Sohn and Park (2013). This metric is included in the revised manuscript (see metric nr. 8 in Section 2.2, lines 102-109). We have searched the peer-reviewed literature for other metrics and have not found any, so to our knowledge, these are the only metrics that have been used in the past.

It is unclear how to quantify the relative relevance of metrics if we do not have a reference (or generally agreed) HC strength metric to which various other metrics can be compared. It is in this context that we introduced a new, energy-based integral metric of the HC strength. The unbalanced energy of the zonal mean circulation is straightforward to derive for gridded datasets and it includes all 3 spatial dimensions of the unbalanced circulation. We are not sure what exactly the reviewer finds confusing about the new energy metric, but the metric is discussed in more detail in the revised manuscript. It is different from the stream function, but also from the omega-based and velocity-potential-based metrics, and more research can be done to refine it, especially to differentiate between the northern and southern branches of the Hadley cell. Our figures A2 and A3 suggest that the global unbalanced circulation is an adequate description of the HC, implying that the associated total energy is an adequate description of the HC strength. We believe that it is a suitable metric also for an intercomparison of reanalyses and climate models analyzed in terms of the normal-mode functions.

Davis, N., & Birner, T. (2017). On the Discrepancies in Tropical Belt Expansion between Reanalyses and Climate Models and among Tropical Belt Width Metrics, Journal of Climate, 30(4), 1211-1231.

Solomon, A., Polvani, L. M., Waugh, D. W., and Davis, S. M. (2016), Contrasting upper and lower atmospheric metrics of tropical expansion in the Southern Hemisphere, Geophys. Res. Lett., 43, 10,496– 10,503, doi:10.1002/2016GL070917.

Sohn, B. J., and Park, S.-C. (2010), Strengthened tropical circulations in past three decades inferred from water vapor transport, J. Geophys. Res., 115, D15112, doi:10.1029/2009JD013713.

Waugh, D. W., Grise, K. M., Seviour, W. J. M., Davis, S. M., Davis, N., Adam, O., Son, S.-W., Simpson, I. R., Staten, P. W., Maycock, A. C., Ummenhofer, C. C., Birner, T., & Ming, A. (2018). Revisiting the Relationship among Metrics of Tropical Expansion, Journal of Climate, 31(18), 7565-7581.

Specific comments:

I find it strange to choose 2 versions of the ECMWF reanalysis instead of 2 new-generation products such as ERA5 and CFSR for a more independent comparison. It's been reported that ERA5 is an improved version of ERAI with many significant fixed errors therefore the discrepancies found by the authors may be attributed to those improvements.

In their specific comment, the Reviewer points out that ERA5 is a more advanced and therefore more reliable reanalysis dataset than ERA-Interim. We could not agree more and we emphasize this out in the revised paper (lines 68-76). Although relatively few evaluations of the CFSR have been conducted and thus its performance is not well-known, we believe that CFSR is much more advanced than the NCEP-NCAR reanalyses. Yet, many researchers would argue that even the NCEP-NCAR reanalysis suffices for the description of the large-scale circulation. Even

though ERA5 is available, many scientists rely on ERA-Interim, and precisely a comparison of tropical aspects in the ERA5 and ERA-Interim, which have been the subject of several recent papers, motivated our study initially. Even though the tropics remain the region with the largest analysis uncertainties (e.g. Žagar et al., 2020, J. Clim), the four modern reanalyses (ERA5, ERA-Interim, JRA55, and MERRA) agree relatively well regarding the large-scale tropical circulation. The total energy of the zonal-mean unbalanced flow shows positive trends in both ERA5 and ERA-Interim, although weaker in the former, further suggesting consistency between the two reanalyses.

Note, however, that the only aim of choosing another reanalysis besides ERA5 was to show that the sensitivity of HC strength trends to the choice of the metric is not an isolated feature of a particular (e.g. ERA5) reanalysis, as stated in lines 299-300. A detailed comparison of the HC strength trends in various reanalyses and the sources of their differences are beyond the scope of this study.

Žagar, N., Zaplotnik, Ž., & Karami, K. (2020). Atmospheric Subseasonal Variability and Circulation Regimes: Spectra, Trends, and Uncertainties, Journal of Climate, 33(21), 9375-9390.

Technical corrections:

L46: suggest replace '…are the trend… the pressure level.' by '…the trend… the pressure level are.'
Corrected as suggested.

Section 2.2: this should go in the result section, not in the methods section
We followed the Reviewers suggestion.

L177: suggest remove 'also'
Corrected.

L182: what do the authors mean by 'merely showcase'?
We have corrected it to a more neutral form: “Figs. 2, A3 reflect the stronger year-to-year variability of seasonal means (compared with year-to-year variability of annual means),…”