Weather Clim. Dynam. Discuss., referee comment RC2
https://doi.org/10.5194/wcd-2021-50-RC2, 2021

Comment on wcd-2021-50

Anonymous Referee #2

Referee comment on "Metrics of the Hadley circulation strength and associated circulation trends" by Matic Pikovnik et al., Weather Clim. Dynam. Discuss., https://doi.org/10.5194/wcd-2021-50-RC2, 2021

It's nice to see Pikovnik et al.'s work comparing indices of Hadley cell strength. As a member of recent working groups on Hadley cell width, I think a paper like this is long overdue. I appreciate Pikovnik's diligence in comparing indices as a function of level and latitude, and noting the impact of changing altitude or latitude on the trends from indices where the maximum value (i.e. of the stream function) is picked wherever it occurs. (An example of a paper that uses this metric is below.)

I am interested in the normal mode-based metric. I found the description in the paper inadequate; wording like "unbalanced flow" is foreign to those of us who haven't read up on the normal mode decomposition. In fact, the relationship between normal modes and the Hadley cell is not intuitively obvious; when I first read that there was a normal mode-based metric, I thought, "this must be a usage of normal mode besides what I'm thinking." But when I read a bit of Zagar & J. Tribbia (2020), I saw that the usage of normal modes was precisely what I was used to, and "unbalanced flow" describes flow that is not in thermal wind balance. Not surprisingly, the overturning one can deduce from the unbalanced flow – see Figures 12 and 14 of Zagar and Tribbia (2020) – looks very much like the overturning one can deduce from the divergent flow – see Figure 2 of Staten et al. (2019).

This regionality may help explain the shortcomings of metric 8, and its lack of correlation with the other metrics. I noticed that metric 8 uses only the energy from the k=0 unbalanced flow mode. This makes sense inasmuch as the Hadley circulation is zonally uniform. But in reality, the Hadley cell has centers of action, and there are even longitudes with counter-Hadley-cell-wise circulations. Much of the activity of interest when we study the Hadley circulation may be regionally focus. Perhaps the observed regional variations that are changing the zonal ψ-based metric are simply not captured by the changing energy of the k=0 unbalanced mode. I wonder whether including the energy from modes with k<=4 (in the tropics and subtropics) into a metric like metric 8 would help to capture the kinds of variations that are producing the temporal variations in ψ-based metrics.

It must be a little disappointing to create such an interesting and seemingly holistic metric…only for it to apparently underperform. It is a credit to the authors that they acknowledge its underperformance in this case. I can't help but wonder if there's more to it, though.

Another outgrowth of Hadley cell width working group efforts that may be of use for the authors was the TropD software written by Ori Adam (see Adam et al., 2018). This software is built for detecting just the kind of metrics Pikovnik et al. use, and does so in a careful manner. I recommend trying the code and verifying that the metrics (or a subset of them) calculated by Pikovnik match those by Adam's software. The detection of extrema is particularly thoughtfully handled in TropD. It would also ensure that your results are intercomparible with those of other recent papers, such as Menzel & Waugh (2019).

Menzel & Waugh's work is worth citing, I think, even though it is nominally about the subtropical jet, rather than Hadley cell intensity. In it, they find that the subtropical jet position is more closely related to Hadley cell intensity than Hadley cell width (and that Hadley cell width is more closely related to the speed of the subtropical jet than to its position). I think this could be cited in the introduction section, as a bit of motivation for studying Hadley cell intensity, as well as another example of a paper that uses metric 1. Also, it is interesting to note that the intensity of the zonal mean overturning (which is mathematically equivalent, I think, to the zonal mean of the overturning due to the "unbalanced" flow) would be so highly correlated with the subtropical jet, which is thought of as being in thermal wind balance.

I am concerned about the use of the Mann-Kendall test for significance. This test requires data to have no autocorrelation. The authors say they use a "modified" version, but the paper does not say how or why it was modified, nor does it address the issue of auto-correlation in the time series. Figure 1 has large regions of weak signal (for example, during March in the lower troposphere over the SH) that are marked as significant. Much of this region is not marked as significant in ERA-Interim (Figure A1). Of course, differences in the trend from one dataset to another do not imply that the trend in one dataset is not significantly significant *in that dataset*. But perhaps the calculation of significance needs to be handled more circumspectly. That said, I appreciated how careful the authors were to distinguish between a trend that happens to be in the time series and trend related to forced climate change.

Little grammatical errors, odd wording, and writing problems are scattered through the paper. I'll mention just a couple, but the paper deserves some careful editing to fix these problems.

The rest of my suggestions are mainly editorial.

- Line 28: “minimum pressure velocity” is mathematically correct but also a bit confusing when what is being described is maximum ascent
- Line 30: “the studies” is better as “studies” here, since a specific group of studies has not been referred to.
- Line 42: I can see why this sentence about Nguyen et al. appears in the same paragraph as the preceding list, but the topic of the sentence feels like it belongs in a new paragraph. It’s also a little disappointing that the results from Nguyen et al. aren’t described here; all that is said is that they have relevant results. I suggest making this a new paragraph, and continuing by mentioning their relevant results. (And if they’re not really relevant, the sentence could perhaps be deleted.)
- Line 46: “we assess how sensitive are the trends” is awkward. “we assess how sensitive the trends are” is better.
- Line 50: It is taken for granted here that readers will understand what the “unbalanced” global circulation is, even though this is not a phrase
- Line 60: While details are fine to be left in citations, some description or summary for the reader would be appreciated
- Line 69: “an extend” should be “the extent”
- Line 70: “a part of the 40-year trends in the HC strength may be due to the multi-decadal variability.” This is (virtually) given. Any time series can be decomposed into a trend + variability + error/noise. Say something more specific, or get rid of the sentence.
- Line 74: I don’t know if “feature” is the right word. It sounds like you are going to describe a particular anomaly. It might be better to say, simply, “Note that trends in…”
- Line 81–82: as I mentioned above, in addition to the reliability of some climate model trends being called into question, I would question the reliability of the statistical significance test used here.
- Lines 145–147: Of course the same feature (namely climatology) can be seen in Figure 1: Figure 1 is the climatology. Why not say, “Normalization accounts for some of these differences; normalized results are discussed near the end of Section 3.2 (see Figure 4).”
- Line 164: I don’t know that “spurious” is the right word to describe point-wise trends. Maybe “non-representative” or “isolated.” The point is not that the trends are wrong, but that they are probably not what the authors or readers are interested in.
- Line 175: “the measure of average HC strength” is vague. If it is a monthly average, any of these ψ-based metrics could be described this way.
- Line 243: “Insignificant correlations are not surprising as this index is largely different from all other indices.” This is not the most satisfying or meaningful sentence. What does largely different mean? If the unbalanced flow is thought to be largely related to Hadley cell-wise circulation, isn’t a higher correlation expected? I can’t imagine the authors went through the effort to define this metric if they expected the correlations to be insignificant.
- Line 246: “unimportant for the overall signal, but it may be important for the trends”. What is the difference between the signal and the trends? Often “signal” is used to describe “trends” and “noise”, is used to describe year-to-year or decade-to-decade “variability.” Did the authors mean “signal” or “climatology”?
- Lines 250–252: Does this result imply that metric 8 is going to be sensitive to increasing tropopause height?
- Line 265: “This was made evident by a new HC strength measure.” It is not clear what was made evident, or how it was made evident.
- Line 285: Based on my working group experience on Hadley cell width, I am skeptical of the idea of a “unified index.” Different metrics for Hadley cell width are of interest to different people for different reasons, even if their trends differ. If we created a unified index, there would be information missing from that index for each group. A unified index for Hadley cell strength would not capture simultaneously the differences between hemispheres, between regions, and in the deep tropical upwelling. Upper tropospheric circulation may be of interest to people studying the stratosphere or the

tropical tropopause region, while mass is concentrated in the lower troposphere, so mass-weighted measures will leave them out.

References:

Adam et al., 2018: dx.doi.org/10.5194/gmd-11-4339-2018

Menzel & Waugh, 2019: dx.doi.org/10.1029/2019GL083345

Staten et al., 2019: dx.doi.org/10.1029/2018JD030100