Review of 'Attributing the occurrence and intensity of extreme events with the flow analogues method'

In this paper, the authors tackle the problem of attributing extreme events to climate change. They take the angle of conditioning extreme events on the regional circulation, as captured by flow analogues, over Europe in the considered cases. This has the advantage of allowing a more 'apples-to-apples' comparison of dynamically similar events, in contrast to the more standard GEV approach based on annual block maxima.
The paper is methodologically focused, and aims to demonstrate the validity and usefulness of this flow-dependent approach on three high profile European extreme events from recent years, and compares the findings to both an alternate detrending procedure and a version of the GEV/ETV approach. They also explore the sensitivity of their method to parameter variations. I think the approach is very interesting; it is definitely sensible and insightful to integrate circulation information into the attribution process.

However the conclusions the authors' analysis provides for the given cases are rather unusual, and after a few careful reads I think the authors either need to better evidence the plausibility of their findings or revise their method. Therefore I am suggesting major revisions. I detail my objection immediately below, followed by a number of minor and typographical comments.

To quickly summarise what the analogue-based attribution shows for each case study:

- The heatwave was made ~4 degrees warmer by climate change, and ~100,000x more likely. The EVT approach claims a similar intensity change (>3 degrees), but only deems the event 10-100x more likely.
- No significant change in probability ratio or intensity of the wind extreme. The EVT approach seems to indicate reduced probability/intensity, but not significantly.
- The precipitation event was made ~100x less likely and was made 8-20 mm/day weaker (or 20-50mm/day weaker; there is an inconsistency between text and figure 3i). The EVT approach suggests a slight, not very significant, increase in probability and intensity.

The precipitation case gives me the most cause for concern. Looking at how the logarithmic fit in 3c,f squeezes the distirbution so drastically (transforming a >25mm/day event into a <10mm/day event in the current climate), it seems very hard to believe this is a sensible transformation to apply. Under the EVT approach in fig 6f meanwhile, the change in distribution is small. To me this makes sense given the fairly weak apparent correlation between RMST and total precip in 3c.
Coming from another perspective, and thinking about the meteorology, the 4 Oct 2021 event was a result of an extratropical cyclone tracking across the Mediterranean, with the anticyclonic anomaly to the east helping to stall its eastward movement. Do we really expect the probability of extreme rainfall when a cyclone hits the coast to have decreased by 90-99% relative to 1950? Or indeed, to have had its intensity ~halved? This runs counter to basic theory (Clausius-Clapeyron relation) and previous work that find little observational change in cyclone properties.
It seems to me that either the logarithmic method of trend fitting is producing spurious results, and/or the similarity metric used to define analogues is not faithful to the driving meteorology (e.g. the analogues don't include cyclones when the true flow does). I suspect that defining analogues based solely on Z500 over a large area is too coarse an approach.

I am also a bit worried about the high probability ratios estimated for the heat case. You mention later that probability ratios have many issues, but the fact remains that your method produces PR estimates 1000x that of the EVT approach. I don't think this is unrelated from the fact that the EVT distribution is positive skewed. While you've fitted a skew normal for the analogues, the distribution has ended up very gaussian so the estimation of tail probabalities is going to be very unreliable. Again, there could be a possible issue with the analogues: it wouldn't take a large distortion of the flow field to put central france under a trough.

I ask the authors to revise or more fully justify these more implausibel findings I've mentioned. It might aid interpretation to actually plot ~4 analogues (say the 15th, 30th, 45th and 60th closest analogues) for each case in the appendix, to see if a mismatch of meteorology between the event and analogues does indeed explain this.

Minor/typographical comments

- Line 12: what is a 'climatological event'? A little unclear to me. Also 'negatively affect' reads more naturally than 'affect negatively'
- Line 15: 'which aim ~~is~~ to'
- Line 31: Given 'very low' doesn't have a well defined quantitative meaning, I'd just say 'estimating low probabilities'
- Line 48: 'These methods both condition the ...'
- Line 57: I think by this point in the introduction it would be good to explicitly state the assumption of your work that changes in the synoptic dynamics themselves due to climate change are negligible/ should be treated separately. The conditional attribution is informative but by definition only a partial attribution to thermodynamic changes. As far as I can see this is not written in the introduction, and should be clearly emphasised.
- Line 62: 'This method is used, for example, by the...'
- Line 68: 'Which lead to exceptional heat in...'
- Line 69: 'Western ~~of~~ Europe'
- Line 71: 'precipitation~~s~~' (and several other times in the document)
- Line 81: Can you motivate why you take a 5 day mean for precip? I presume to focus on the synoptic driving.
- Line 86: Even 1 degree is a fairly strong constraint on shifts. It might be interesting to see results for a 2.5/3 deg box average centred on the three points examined in figure 3.
- Line 103: Clearer to say "thickening due to warming of the atmosphere".
- Line 111: Its not totally clear in what space you compute the analogues. The full high-dimensinal space of gridpoint values within each region? That's giving you a O(1000) dimensional space. Given that distance vectors in high-dimensional space tend towards equal length (https://link.springer.com/chapter/10.1007/3-540-44503-X_27) which may lead to poor ability to detect analogues, did you consider computing analogues in an O(10-100) dimensional PCA space?
- Line 127: impossible from the thermodynamic perspective, but this could be tackled by considering how synoptic conditions are being impacted by climate change.
- Line 136: What is the motivation for using surface temperature for something like precip, given condensation happens aloft? If RMST is used just for simplicity then maybe make a comment about this somewhere.
- Line 199: Worst-> worse
- Fig 3abc: You should state in the figure caption that the 'detrended' data is shifted to the 2020 RMST value. It makes sense for your application, but isn't a standard detrending, where you'd expect to end up with zero mean displacement of the data.
- Line 263: The intensity change your report does not correspond to what you show in fig 3i.
- The captions for figures B5 and B6 are the same.
- Line 345: missing reference.