Weather Clim. Dynam. Discuss., referee comment RC2
https://doi.org/10.5194/wcd-2021-1-RC2, 2021

Comment on wcd-2021-1

Anonymous Referee #2

Referee comment on "An unsupervised learning approach to identifying blocking events:the case of European summer" by Carl Thomas et al., Weather Clim. Dynam. Discuss., https://doi.org/10.5194/wcd-2021-1-RC2, 2021

In this paper, the authors introduce a new blocking index based on self organizing maps. They compare this index to other commonly used blocking indices and to a ground truth time series they build based on expert judgement. They show that this new index generally performs better than other indices. I find this paper very well written and structured, with a very thorough sensitivity testing of their index to several key parameters. I have a few comments and questions that I list below. I recommend to accept this paper after some revisions.

General comments:
You say your GTD is objective, but you make some heuristic choices too (namely the region and time period), so it's not completely objective. It's also a little bit unclear to me how you label your blocking periods from the maps shown in Figures 1 and 2. Do you manually look at each group of maps and identify blocking highs?

Could you make it clearer in the text what's the added value of the SOM-BI compared to the GTD? As it is presented, the GTD gives supposedly better results than the SOM-BI so why do you need to build the SOM-BI? This may be easier to understand once the methodology to construct the GTD is clearer.

The sinuosity index seems a bit disconnected from the other indices since it's not a direct measure of blockings. I would advise to remove this part of the analysis from the article, especially since it's a hemispheric measure and your study applies to regional blocking indices.

I am not a specialist of SOMs but I am more used to k-means algorithm and weather regimes. You never cite this part of the literature but I think it would be important to do so since that's something familiar for a large part of the atmospheric dynamics community.
Michelangeli, Paul-Antoine, Robert Vautard, and Bernard Legras. "Weather regimes: Recurrence and quasi stationarity." Journal of Atmospheric Sciences 52.8 (1995): 1237-1256.
Vautard, Robert. "Multiple weather regimes over the North Atlantic: Analysis of precursors and successors." Monthly weather review 118.10 (1990): 2056-2081.
It would be good to discuss the differences between your approach and what is done in

the context of weather regimes.

Specific comments:
The abstract is a bit misleading. The SOM-BI doesn’t work well with the 2019 case and this should be acknowledged.

L. 58 do you mean objective or subjective?

L.254 I think there’s a mistake with R=0.03. Shouldn’t it increase compared with 0.19?

L.263 remove « be »

L.309 rephrase « are not more generally »

Figure 7 and 8 look very flat to me. Since those are the curves you use to determine your optimal number of nodes I would discuss this in the paper and potentially introduce some sort of confidence interval around your optimal k.

L.419 and 421 I think you’re talking about Figure 10 here, not Figure 9.