Earth System Dynamics Discussions

Earth Syst. Dynam. Discuss., referee comment RC2
https://doi.org/10.5194/esd-2022-29-RC2, 2022

Comment on esd-2022-29

Anonymous Referee #2

Referee comment on "Classification of synoptic circulation patterns with a two-stage clustering algorithm using the structural similarity index metric (SSIM)" by Kristina Winderlich et al., Earth Syst. Dynam. Discuss., https://doi.org/10.5194/esd-2022-29-RC2, 2022

Review of the manuscript entitled: "Classification of synoptic circulation patterns with a two-stage clustering algorithm using the structural similarity index metric (SSIM)" by Kristina Winderlich, Clementine Dalelane and Andreas Walter

Summary

The authors develop a new classification method for synoptic circulation patterns with the aim to extend the evaluation routine for climate simulations. Its unique novelty is the use of the structural similarity index metric (SSIM) instead of traditional distance metrics for cluster building. This classification method combines two classical clustering algorithms used iteratively, hierarchical agglomerative clustering (HAC) and k-medoids. The authors apply the classification method to ERA-interim and NCEP1 reanalysis, and CMIP6 models. The authors wish to demonstrate that the built classes are consistent, well separated, spatially and temporally stable, and physically meaningful. Finally, the authors rank the CMIP6 models according to their ability to represent the weather types using different quality indices.

Dear authors,

The purpose of using synoptic circulation patterns to evaluate climate models is a welcomed aim, but is not the first time this is done, as it may seem from the text. Indeed, the ability of models to capture the characteristics of synoptic patterns is an important aspect of improving climate model simulations. The SSIM is generally an interesting and

seems to be promising approach for the classification of weather regimes. The article is generally well written, however it should be extended to serve as a high quality research article in ESD.

My comments and suggestions to improve the manuscript are as follows:

General comments

- Many classification algorithms attempt to categorize weather types/regimes over the Atlantic-European-Mediterranean region. If the authors suggest a new procedure, they should at least demonstrate why their classification is better than other classification procedures. Indeed, the authors try to explain their choices, but do not demonstrate how their procedure is superior in comparison to other classifications. Perhaps the authors can randomly select days and subjectively see for how many of them the classification does a decent job? Comparing to the original classification you mention in the text would then provide a semi-quantitative way of demonstrating the improvement from one classification to the other.
- Forty-three classes seems a rather large number of weather types and can probably be significantly reduced by some sort of EOF analysis. If not, it should at least be explained why the authors do not use this approach as it is very common. Furthermore, I would like to see some further explanation on how do these synoptic types relate to the four canonical weather regimes.
- The CMIP6 model evaluation section in its current form is rather short and does not provide very useful information for model developers. This section should probably be extended. It would be nice to have some discussion as to why you think some models are better or worse. Additional analysis is of course welcomed, but should probably be balanced with the length of the article.

Specific comments

Abstract

- What do you mean with physically meaningful? There may be different meanings to physical, and you should probably clarify this in the text.
- Line 10: This sentence should be at the very end of the abstract.
- Do you think your classification would be useful for extended-range weather forecasts? If so, mention this and in the abstract and discuss in the conclusions.

Introduction

- Line 43 – 47: From the introduction, it sounds as if you are the first and only group evaluating models based on weather regimes. However, there is an increasing body of knowledge working in this direction. To name a few articles:

References

Dorrington, J., Strommen, K., and Fabiano, F.: Quantifying climate model representation of the wintertime Euro-Atlantic circulation using geopotential-jet regimes, Weather Clim. Dynam., 3, 505–533, https://doi.org/10.5194/wcd-3-505-2022, 2022.

Fabiano, F., Christensen, H.M., Strommen, K. et al. Euro-Atlantic weather Regimes in the PRIMAVERA coupled climate simulations: impact of resolution and mean state biases on model performance. Clim Dyn 54, 5031–5048 (2020). https://doi.org/10.1007/s00382-020-05271-w

Hochman A, Alpert P, Harpaz T, Saaroni H, Messori G. 2019. A new dynamical systems perspective on atmospheric predictability: eastern Mediterranean weather regimes as a case study. *Science Advances* 5: eaau0936. https://doi.org/10.1126/sciadv.aau0936

- Line 58: Please discuss the number of regimes some more. There are a few articles focusing on this aspect in the literature. Some use two regimes (Wallace and Gutzler, 1981), others use four (Vautard 1990), six (Falkena et al., 2020) or seven (Grams et al., 2017) regimes. This is important as you use an outstanding number of 43.

References

Falkena, S. K., de Wiljes, J., Weisheimer, A., & Shepherd, T. G. (2020). Revisiting the identification of wintertime atmospheric circula-tion regimes in the Euro-Atlantic sector. Quarterly Journal of the Royal Meteorological Society, 146, 2801–2814. https://doi.org/10.1002/qj.3818

Grams, C. M., Beerli, R., Pfenninger, S., Staffell, I., & Wernli, H. (2017). Balancing Europe's wind-power output through spatial deployment informed by weather regimes. Nature Climate Change, 7, 557–562. https://doi.org/10.1038/nclimate3338

Vautard, R. (1990). Multiple weather regimes over the North Atlantic: Analysis of precursors and successors. Monthly Weather Review, 118,2056–2081. https://doi.org/10.1175/1520-0493(1990)118<2056:MWROTN>2.0.CO;2

Wallace, J. M., & Gutzler, D. S. (1981). Teleconnections in the geopotential height field during the Northern Hemisphere winter. MonthlyWeather Review, 109, 784–812. https://doi.org/10.1175/1520-0493(1981)109<0784:TITGHF>2.0.CO;2

- Line 64-66: This is a very strong critic on all prior classifications and should be further explained why none fit your purpose. These classification procedures were all used extensively in the literature. If you state this, you should at least demonstrate how your classification is superior.

Data and methods

- Line 80: If you use ERA-interim and not ERA5 reanalysis, you should at least say why, and mention some of the studies comparing the two data sets. I do not expect much difference for large-scale weather regimes, but this should be at least discussed.
- Line 82: Please justify why you use 12:00UTC and not daily or all 6-hourly data.
- Line 82: How did you coarse grain the data and why to 2×3 degrees?
- You often use 'synoptic scale', but I think it is more accurate to consider these regimes as large-scale features. I would try being more accurate on this. Perhaps change throughout the text.
- Line 95: Why 151 days of smoothing? Please justify this choice.

Results

- Lines 436-440: I do not completely understand how you obtained high resolution relative to coarse resolution in figure 9.
- Line 454-456: Your motivation was not to use centroids in the introduction and methods section, but then you test your medoids and say that they are very similar to the centroids. Is this not a circular argument?
- Section 4.6: Perhaps provide some illustrations of the different classes in the CMIP6 models, in addition the quality indices in the table.
- Table 3: I believe that there is not much difference between the models in the 'transit' and 'persist' values because there are so many classes. In addition, for the other indices the standard deviation is rather low, which is a bit surprising for more than 30 models. They all do pretty much the same job, which is again a bit surprising.
- Are the models evaluation criteria significantly different from one another? I think you should test this.

Conclusions

- This section is rather very short and should have a bit more discussion with respect to other articles evaluating models using a classification procedure. The article would also benefit from explaining what is better or similar in the new classification with respect to other methodologies used in the literature. The potential use of this methodology in climate projections or extended-range weather forecasts should probably also be discussed.

Technical comments:

- Line 82-84: Please rephrase, something is missing here.
- Line 307: This should be 'Results' and not 'Method' section.
- Line 318: Change 'gives us an evidence that' to 'provides evidence that'.
- Line 357: Change 'gives an evidence that' to 'provides evidence that'.

Figures:

- Figure 4: It is very hard to see anything with so many panels.
- Figure 10: I think you mixed up between left and right in the caption. In addition, are there significant difference in the right panels?
- Table 3: It should probably be DJF for winter in the upper row and not 'JDF'.