Review comments on esd-2020-61 "Present and future synoptic circulation patterns associated with cold and snowy spells over Italy", version 30 Nov 2021, by M. D' Errico and co-authors

GENERAL COMMENTS

This is my third review of this manuscript. Again, there have been extensive changes that solve some of the previous problems but also raise new questions and concerns.

I particularly appreciate the condensation of the previous case-by-case results to the cluster-wise analysis in Table 1. The way in which the results are represented in this table is useful, because it shows that the large relative increases in the frequency of the circulation analogues in fact reflect relatively modest increases in the absolute numbers.

As the authors now apply their analogy search to anomalies in Z500, which I presume to have been calculated against the time means in the RCP simulations, the increase in the frequency of good analogies apparently cannot be explained by biases or changes in the time mean state. Yet, the manuscript still leaves it unclear to me why this increase takes place. I realize that it may be difficult to give a fully satisfying answer to this question. However, one way could be to pick up a fixed number (e.g. top 2%) of the best analogies for both the CTRL and the RCP simulations, and calculate the local root-mean-square difference between these and the cluster centroids for each grid point in the analysis domain (22.5º-70ºN, 10ºW-70ºE). Doing this and comparing the resulting maps, it would at least be possible to deduce whether the analogies in the RCP simulations are better in a specific part of the domain or whether the improvement is more or less evenly distributed in the area.

In this version of the manuscript, bias corrections have been introduced that adjust both the time mean bias and the standard deviation. The correction for the mean bias seems redundant as far as the analogies are based on anomaly values only. There is also a risk that the correction of the standard deviation makes more harm than good. If the standard deviation biases in PlaSim are different in different parts of the domain, then the correction of the standard deviation may result in anomaly patterns that differ (e.g., in the locations of their maxima and minima) from the originally simulated patterns. In such a situation, the dynamical interpretation of the results becomes difficult.

What made me pay attention to such potential bias correction artefacts is Fig. 4g-h. These maps show, for the RCP8.5 scenario, Z500 height fields with a strong (> 6000 m?) maximum in the northernmost part of the domain. This is completely different from the patterns in the CTRL (Figs. 4a-b). Moreover, as there is a deep minimum in SLP in the same area in Figs. 6g-h, such a maximum in Z500 would only be consistent with the hydrostatic balance if the temperatures in the northern part of the

domain were much higher than those elsewhere. However, Fig. 7g-h shows that, even for RCP8.5, T850 is lower in the northern than the southern part of the domain. Thus, Z500, SLP and T850 and SLP are physically inconsistent with each other. Such physical inconsistences could result if the correction of standard deviation is also affecting the climate change signal and not only the variability around the mean state.

Therefore, my advice is to repeat the analysis without the standard deviation correction, or at least modify the implementation of this correction so that it has no impact on the time mean state in the RCP simulations.

The second main weakness in the revised manuscript concerns the new Section 4 on the CMIP5 simulations. This seems like a last-minute addition since it is mentioned neither in the abstract nor the concluding section in the manuscript. More importantly, the description of what was done is so brief that it is difficult to understand what the results mean.

On L283-285, it is stated that the (observed?) daily Z500 anomalies for the 32 cold spell events were embedded into the historical simulations for 1951-2000 and the RCP4.5 and RCP8.5 simulations for 2051-2100. Does this mean that you only used the time mean Z500 fields in the CMIP5 simulations, and then generated Z500 fields corresponding to the cold-spell cases by adding the observed anomalies to these time mean fields? This sounds rather different from what you did with PlaSim, for which the analysis focused on a comparison of the simulated and observed anomalies. Or did you in fact use the daily Z500 data in the CMIP5 models and repeat the analysis done for PlaSim (which would require a major effort)?

The follow-up on the two types of bias correction on L291-296 is equally difficult to grasp. First, “bias correction” is a somewhat misleading word for removing the spatially averaged Z500 trend (which I assume to mean the area and time mean difference in Z500 between 2051-2100 and 1951-2000). Second, what does the bias correction do if the spatial mean trend is preserved – do you also correct for the biases in the local 1951-2000 mean values? In either case, it is intractable to me how you constructed the anomalies that were compared with the observed (cluster 1 or cluster 2 mean?) anomalies in the end.

Please add detail to the description of the CMIP5 analysis. Otherwise, the results in Section 4 are impossible to interpret. Also, you should mention this analysis both in the abstract and in Section 5, as it appears to be an important part of your work.

Because of these reasons, the manuscript still requires major revisions. Further comments on the details are given below.

SPECIFIC COMMENTS

1. L6-8. The CMIP5 analysis should also be mentioned in the abstract.
2. L27-40. Consider also citing O'Gorman (2014): https://www.nature.com/articles/nature13625
3. The legend in 1 indicates that the size is proportional to the duration. Does this indicate that the *diameter* of the circles if proportional to the duration, or that their *area* is proportional to the duration?
4. Figures 3a-d. Please include the temperature anomalies in the maps. You can add them as contour plots and simultaneously retain the shading for the absolute values.
5. L175. A better general reference to the RCP scenarios: van Vuuren, D.P., Edmonds, J., Kainuma, M. et al. The representative concentration pathways: an overview. Climatic Change 109, 5 (2011). https://doi.org/10.1007/s10584-011-0148-z
6. L177-178. The actual CO_2 concentrations in the RCP scenarios are lower. The quoted numbers are equivalent CO_2 concentrations, which include the net effect of all anthropogenic greenhouse gas forcing.
7. Table 1, cluster 2, RCP85. +71.3% (in parentheses) appears too small. Should rather be ~121%?
8. Caption of Table 1. Please specify the meaning for the numbers that are given for the RCP scenarios (i) before the parentheses (e.g. 0.978) and (ii) in the parentheses (e.g. 0.977).
9. L283-285. Did I understand this correctly? You simply take the 50-year time mean of Z500 in the CMIP5 simulations and add the observed anomalies in the 32 cold spell cases?
10. L284-285. What was the time resolution of the CMIP5 data? Daily or monthly? Overall, Section 4 is difficult to follow because the methodology is not explained in sufficient detail.
11. L297. Euclidean distance relative to the mean of the cluster (1 or 2) in which the event belongs to?
12. L299. Why does the inclusion of the spatially averaged trend lead to closer analogies? One would expect that, with a general warming of climate, the Z500 anomalies relative to the 1951-2000 mean will become systematically positive. If anything, this should make the analogues worse.
13. Section 5. The CMIP5 analysis should also be mentioned in this section.

TECHNICAL COMMENTS

1. L203. to the control and RCP simulations?
2. L214. persistence?
3. L216. cold spells
4. L254. these results
5. L292. that trend / those trends
6. L297. more slightly or more strongly?

7. L298. these decreases
8. Caption of Figure 10, L1-2. each dot represents
9. Caption of Figure 10, L2. given
10. L324. requirements of the Paris agreement (Arias et al. 2021).