General Comments

The authors analyse the links between the occurrence of anticyclones, dry spells and heat waves in a large ensemble of CMIP5 GCM simulations for the historical period. Additionally, ERA5 and E-OBS data is considered in the evaluation. The authors conclude that the discrepancies between the GCMs identified in terms of the duration of dry spells and extreme temperatures are related with the GCM biases regarding temperature and precipitation themselves. While the topic of the manuscript is surely interesting and I acknowledge that there was a huge amount of data processing involved, there are also severe shortcomings, the main one being the way that "anticyclones" are considered, which is methodological not sound (see major point). Given that this was the only aspect broadly related with atmospheric dynamics (the core of WCD), I'd argue that the manuscript in its present form is also out of scope of the journal. For both reasons, and given that this shortcoming is pretty fundamental to the whole analysis, I'd like to recommend its rejection in the present form (as the whole calculations and analysis would need to be redone).

This say, I'd strongly encourage the authors to take up this task and resubmit the manuscript for further evaluation. If resubmitted to WCD, the aspect of atmospheric dynamics should be strongly strengthened so it fits the scope of the journal (e.g. also links to blocking, synoptic weather types, anticyclone dynamics). If the authors do not which to change the focus of the manuscript so strongly, I'd suggest the resubmission to a different journal – notably NHESS, which focus primarily on impacts. I would be willing to review the paper again upon resubmission.

Main Comment

The main shortcoming in the present study is the way "anticyclones" are dealt with. For me, an anticyclone is a high-pressure centre with clockwise rotating winds and large-scale divergence at the surface. At upper levels, it is typically associated with a cut off low / ridge / blocking system, where upper-level convergence occurs, thus leading to large-scale subsidence in the area of the surface high-pressure centre. As described e.g. in a recent review paper published in WCD (Kautz et al. 2022; https://doi.org/10.5194/wcd-3-305-2022), a persistent anticyclone / blocking over Central Europe in the summer typically leads to heat waves and dry spells collocated with the centre of the system (see their Figure 2b). However, on both flanks of the system you often observe heavy precipitation associated e.g. with moist air intrusions on the western flank of the system (same figure).

Given this, I am really puzzled that the authors "define an anticyclone" as local exceedance of MSLP above 1012 hPa over five days. This value is below the average mean MSLP for a considerable part of Europe (particularly in the summer months!), and even below the global average MSLP value! Given the often quite stationary weather conditions in the summer, five days is also no real constraint. So we are not even looking at above average pressure conditions. And of course, this has nothing to do with atmospheric dynamics and cannot provide any insight either on the exact location of the anticyclone and where different types of extremes may be expected. If a simple metric is needed, I'd use chose something based on MSLP anomalies to the monthly mean (or summer) MSLP fields, thus as an indication of the anomalous circulation associated with the high-pressure anomaly.

This explains why several of the following evaluations/results do not really meet the expectations (e.g. compared to what you'd expect for blocking, e.g. around lines 265) or are not really understandable (at least to me), e.g. Figs 3 and low collocation of dry spells and anticyclones for Southern Europe (as acknowledged by the authors also in the text), and several other figures, particularly Figure 8. While I do like the idea of the paper (which is my I accepted the review), I do thing the methodology is flawed and the presented results are thus unfortunately not sound.

We thank the reviewer for their time in reviewing this manuscript, it is much appreciated. We accept that the way we have treated anticyclonic conditions is fundamentally flawed and so we have updated the analysis and applied an algorithm from Sousa et al. (2021) that detects anticyclonic features, namely sub-tropical ridges and atmospheric blocking using geopotential heights at 500 hPa. We combine these features under the same definition, anticyclonic systems, as both will have the same local effect on precipitation and temperature and both can also occur within the same life cycle of an anticyclone. The updated analysis assesses the frequency of anticyclonic systems in Europe according to the chosen algorithm, their local influence on the persistence of dry spells and the link between model biases in the frequency of anticyclonic systems in models and biases in both dry spell persistence and the magnitude of temperatures during dry spells in models. The text has also been improved to include a greater discussion of the literature surrounding anticyclonic systems, their influence on dry spells and temperature, as well as implications of the results. We feel the analysis and strength of the conclusions around anticyclonic systems is much improved and we hope that the paper will provide a meaningful contribution to the discussion around the importance of assessing the influence of biases in the large-scale circulation on relevant hazards when using climate models to aid future decision making.

Changes have been made to the text throughout and new sections have been added related to the additional analysis of anticyclonic systems.

- **Description of the methodology to detect anticyclonic systems is described in section 3.2 (L126-189).**
- **The new results related to anticyclonic blocking are provided in section 4.5 (Figure 7 and Figure 8).**

Minor comments

a) It is not clear to me why version 16.0 of E-OBS is being used, as we are currently on version 24.0 (see https://www.ecad.eu/download/ensembles/download.php). There have been quite a few important updates since.

The older version was used as this had been downloaded and processed as part of a project that officially finished a few years ago. Having read through the updates, we do not note any major update that would yield this version unusable and do not expect to see large differences in the results from this version and the latest version. We are confident that the results and conclusions are insensitive to this choice.

b) I think it is a very strong statement to say that the “combination of dry spells and extreme temperatures” has not yet been assessed in CMIP5 models. Please weaken the statement.

This statement has been weakened (L66).

c) The description of the results is often not understandable, e.g. the description of figure 3b (likes 228-236). Please enhance.

Thank you for pointing this out, we have enhanced the description of the results throughout.

d) Several of the references are missing page ranges, issue numbers, particularly for AGU journals, please enhance.

Thank you for checking this and we apologies for this mistake. All references have been updated to include this information in format required by the journal.

This manuscript investigates the representation of compound dry and hot spells in Europe in the CMIP5 data set. The model data is compared to EOBS. The results show that CMIP5 models struggle to capture the duration and intensity of these compound events. The manuscript is well written, the figures are clear and the results are relevant. I recommend to publish the manuscript after major revisions as detailed below.

We thank the reviewer for taking the valuable time to review this manuscript, it is much appreciated.

Major points:

The choice of a constant MSLP threshold needs to be further motivated and discussed. There are several issues with this choice. I) heat lows can from over the Iberian peninsula (https://rmets.onlinelibrary.wiley.com/doi/abs/10.1256/qj.01.189) during hot conditions breaking the link between MSLP and high temperatures, while the overall tropospheric circulation is still anticyclonic. Ii) in locations with high orography the correction of the surface pressure to MSLP might introduce biases. Iii) the climatologically lower pressure at higher latitudes leads to longer exceedances over the 1012hPa threshold compared to lower latitudes.

We accept that the way we have treated anticyclonic conditions is flawed and so we have updated the analysis and applied an algorithm from Sousa et al. (2021) that detects anticyclonic features, namely sub-tropical ridges and atmospheric blocking using geopotential heights at 500 hPa. We combine these features under the same definition, anticyclonic systems, as both will have the same local effect on precipitation and temperature and both can also occur within the same life cycle of an anticyclone. The updated analysis assesses the frequency of anticyclonic systems in Europe according to the chosen algorithm, their local influence on the persistence of dry spells and the link between model biases in the frequency of anticyclonic systems in models and biases in both dry spell persistence and the magnitude of temperatures during dry spells in models.

Changes have been made to the text throughout and new sections have been added related to the additional analysis of anticyclonic systems.

- **Description of the methodology to detect anticyclonic systems is described in section 3.2 (L126-189).**
- **The new results related to anticyclonic blocking are provided in section 4.5 (Figure 7 and Figure 8).**

A direct comparison of absolute temperatures between EOBS and CMIP5 (Figure 5) will be strongly affected by the representation of the orography and coast lines within CMIP5. A comparison relative to a percentile might be more meaningful.

We agree that the representation of the coastline and orography may play a role in the temperature biases. However, we do not understand the rational behind why a percentile-based approach would remove these effects as they will be seen across the temperature distribution. However, we have added text (L293-295) to note the influence of coastal effects on the interpretation of the biases in those locations.

Please control for multiple testing in all analyses using the FDR (see Wilks 2916, https://journals.ametsoc.org/view/journals/bams/97/12/bams-d-15-00267.1.xml)

The cited paper refers to parametric tests which require statistical assumptions. Our approach uses non-parametric bootstrapping which does not require such assumptions. In the approach, we randomly shuffle seasons to break the seasonal dependence between the precipitation and temperature series and calculate the metric. This is repeated 1000 times and provides an indication of whether the result can be achieved by random chance. As the approach is non-parametric, FDR does not apply here, and we interpret the presence of stippling as there being a < 5% chance of the result occurring by random chance.

How relevant is the representation of summer convection in the models for the duration of the dry spells?

This is an interesting question though one which we cannot answer here. It would likely require a detailed analysis of the models with specific types on convection parameterisation schemes and/or a comparison with a high-resolution convection permitting climate simulations.

Minor points

Abstract: long-duration ï sub-seasonal (Long duration is per se not very clear, it could also refer to spells that last for several years)

Thank you for the suggestion, this has been added to the abstract.

38 Zscheischler 2020/2021 is missing in the list of references

We apologise for this omission; we have updated the reference list.

Add Ridder et al. 2022 to the list of references https://www.nature.com/articles/s41612-021-00224-4

Thank you for highlighting this paper, we have now cited this in the introduction (L45).

96 Is the mean taken across all spells? The definition is not yet fully clear.

Yes, we calculate the maximum temperature within each dry spell lasting longer than 5 days. Tx_{DS} is then the average maximum temperature from all those dry spells.

182 IPPC ï IPCC

Thank you for noticing this mistake, it has been corrected.

410ff Include the results of Zscheischler and Seneviratne (https://www.science.org/doi/full/10.1126/sciadv.17002639) in the discussion.

This has been included in the discussion (Line 537).

Figure 1a I recommend to use a colormap with only one color, two colors suggest a change in sign.

Thank you for this suggestion, and we agree that it might appear as a change in sign though as the paper does not assess changes in the hazards, we feel this is not a large issue. The colormap has been chosen to remain consistent with a previous paper (Manning et al., 2019), and we chose this colormap as it highlights the large difference between dry spell lengths in Northern and Southern Europe. As such, we prefer to keep the current colormap.

Panels c,d,e in Figure 3 do not fit the description and look the same as panels c,d,e in Figure 7, there may have been a mix-up.

Thank you for pointing this out and we apologies for this mix up. This mistake has been corrected.