REVIEWER1

General comments

In the revised version of this manuscript, the authors have addressed the major issue raised in the first review with respect to the bias correction of the duration of dry spells. Instead of applying quantile mapping directly to the duration pdf, they have applied a bias correction to the underlying time series before calculating duration. Additionally, they have also included an additional assessment of two multivariate bias correction methods. In their analysis, they assess changes in the duration of dry spells and temperatures during dry spells over the Pyrenees in future projections from an ensemble of regional climate models. Specifically, they assess changes in the annual mean duration of dry spells that exceed the annual 95th percentile of duration as well as changes in the annual mean extreme magnitude of dry spells, which is the annual mean of temperatures that exceed the 95th percentile of temperature during dry spells. They find that extreme temperatures generally increase during dry spells in future projections, while the changes in duration can vary depending on the assessed region. The analysis is clean and the results are well presented. There is also novelty in the presentation of changes in the bivariate distribution, which is a nice a feature of this paper. However, some clarifications are required in the text, particularly with respect to the event definition which I found created confusion when interpreting results. Furthermore, it would also help if some discussion was added on the model biases and the use of bias correction. With these minor changes, I would recommend this manuscript for publication.

"We gratefully acknowledge the reviewer's comments and the revision of our manuscript again. The article has been revised in accordance with the referee's comments and suggestions, which are addressed below. Our responses appear in italic and quotation marks."

Comments (P: page, L: Line)

P7 L161: I think the explanation of M and EM could be simplified as I found the description of EM and M as 'events' confusing. From what I understand, the event is the dry spell and this is has duration D which is a characteristic of the event. Then, M is the conditional distribution of temperatures during dry spells while EM is the conditional distribution of temperature during dry spells that exceed the 95th percentile of temperature. I believe changing the description of the calculation of the variables would make things easier for the reader. I would recommend removing any reference to M or EM as events.

The authors agree with the reviewer's comment. In fact, it is an aspect that two of the reviewers emphasize that should be improved. For this reason, we have rephrased L161-166 and changed figure 2 to a real case in two different years.

P7 L175-176: Why do you take the time series from one grid cell? Would it not be more consistent with observations if the mean of all grid cells in the region was taken?

The EURO-CORDEX data have a resolution of ~11 km, while the starting grid has a resolution of 1 km. The regional averaging of the observational data is intended to smooth the observational data slightly, and thus avoid problems already known when using bias correction methods for downscaling (inflation problems in the corrected series -Maraun, 2013- and inability to generate daily subgrid variability).

P8 L185-192: How do you treat zero values in this quantile mapping? Are the lowest precipitation wet days in the model simply converted to dry days?

To correct the drizzle effect in the QM, we use the threshold of 1mm/day. We have added this

information in L200-201.

P9 L211: Could you clarify what is meant by 'annually aggregated data'?
We have rephrased it. Please see L220

P9 L213: I think there is a comma missing after 'which was bias corrected'.

Right, comma added.

P16 L195-315: As is highlighted here by the authors, the results shown in Figure 8 show that the corrected CDFs are in many cases very similar to the uncorrected version or worse. Is there a value in using the 'corrected' dry spells vs. 'uncorrected'?

The results show trivial improvements, as they transfer the intrinsic error of the model to the bias correction performance. Precisely, we have mentioned between lines 316 and 318 that the results should be taken with caution. We have discussed these results in the discussion section.

P18 Figure 9: I think it would be informative to include a QQ plot of the uncorrected temperature also, to give an idea of the biases present in the temperature series.

The authors are grateful for the reviewer's suggestion. We have included the QQ plot of the uncorrected temperature in the supplement (Fig. S4). Figure 9 of the manuscript allows to see the performance of the UBC and MBCn at the most extreme values.

P19 L327: Could you clarify what is meant by 'magnitude of intervention' in the section title?

This section title was quite confusing. We changed by: Future changes in the variables underlying the compound event

P19 L329-335: Are these results shown for 'uncorrected' or 'corrected' time series? Which bias correction method is selected if the corrected time series are assessed?

The reviewer is right. The authors forgot to mention that the BC method used to show the projected results was the MBCn. Please see L338-339

P21 Figure 11: I assume this figure shows the multi-model mean projected change? If so, it would helpful to state this in the caption.

We have stated the reviewer suggestion in this caption.

P22 Figure 12: I think this figure still needs some clarification. Does each point in the scatter plot represent the multi-model annual mean of D and EM in a given year? The use of the notation D and EM in the caption is also quite confusing here. On P6 in the event definition, this notation is used to represent individual events or days (for M and EM), but here the notation is used to refer to annual mean anomalies of D and EM. This is also the case in Figure 10 and 11. I would suggest using different notation for individual events and annual mean values of those events.

Yes, each point in the scatter plot represents the multi-model annual mean of D and EM in a given year. We have rephrased this figure caption, as well as the respective figure caption in the supplementary material, following your suggestions.

Discussion section: The use of bias correction to correct the distribution of dry spells will simply take the least wet days and convert them to dry days. As noted in Maraun et al. (2017), this may correct biases resulting from the drizzle effect but not biases resulting from topographical issues or underestimation in the persistence of anti-cyclonic conditions. I think the authors should add some

discussion on this point as well as add some discussion on the results of their bias correction analysis and the performance of climate models in their representation of dry spell both before and after bias correction.

The authors have added a discussion paragraph on the performance of bias correction in resolving bias in dry spells (temporal dependence). Please see lines 422-427.

Reviewer 2

General comments

Thank you for the opportunity to review the revised version of the manuscript "Assessing internal changes in the future structure of Dry-Hot compound events. The case of the Pyrenees". The authors had put in considerable effort in addressing most of the aspects raised by the three reviewers, particularly in terms of the regionalization procedure, the event definition, and the multivariate bias correction. During the review process the authors have also found that they were estimating D, E and EM events in Figs. 3 and 4 for the full period 1981-2015 and not annually, and hence, only Figure 1 has not been changed from the originally submitted manuscript. In addition, the authors have added to the revised manuscript a new figure (revised version Figure 5), that was not requested by any of the reviewers, but it is my view that providing an illustration of the large-scale drivers of the compound event in analysis is an added value to the manuscript and that it supports the results. In general, I consider that most of the reviewers' comments were addressed and that the key necessary changes were performed to the manuscript, greatly improving its original version. However, I still have minor comments to make, which I outline with more detail below, and I leave the decision on these final suggestions to the journal editor, if I may.

"We gratefully acknowledge the reviewer's comments and the revision of our manuscript again. The article has been revised in accordance with the referee's comments and suggestions, which are addressed below. Our responses appear in italic and quotation marks"

Specific comments

Figure 2 - In the reply to my suggestion to improve Figure 2 with an illustration of an observational year, the authors considered that the scheme is enough to illustrate the procedure. However, I still think that it was a common point to all the reviewer's comments that the definition of the event required clarification. In addition, as an answer to a Reviewer's comment the authors have exemplified with the particular year 1981 of the grid cell i,j, where there were 2 dry spells accounting for 15 and 20 days, respectively. In this way, we agree that the exemplification with an observational year helps to better understand the procedure, and this would be a good way to improve Figure 2 and the clarification of the event definition.

The authors have considered the reviewer's suggestion. In this sense, an example of the detection of the extreme dry spells (D), the conditional distribution of temperatures during dry spells (M) and the conditional distribution of temperature during dry spells that exceed the 95th percentile of temperature (EM), in a cell i,j has been included in the manuscript as an example. In this sense, figure 2 shows the time series of daily maximum temperature and daily precipitation for the years 1995 and 2006 for the range x = 2.9, y = 42.5, thus illustrating the procedure followed in this work.

The authors could also mention that they performed a sensitivity analysis using the threshold 90th, and briefly outline the main differences in the analysis using a slightly lower threshold.

We have mentioned that we performed a sensitivity test to select the adequate threshold. The selection of the 95th percentile allows us to obtain sufficiently long and robust dry spells in the most humid areas of the Pyrenees. Please see L.161-162

P2 L43 - I'm afraid the authors have misunderstood my suggestion to include a short introduction in one sentence to the definition proposed by Manning et al (2019), as this work is an extension of the definition proposed by Manning et al (2019). I was not suggesting to just removing the parenthesis to the reference, but to guide better the reader of the introduction that

the magnitude of a Dry-Hot event concerns the temperature and that the duration of a Dry-Hot concerns the length of the dry spell, as proposed by Manning et al (2019). This is not an obvious definition, in my point of view, and when reading the introduction for the first time the concept could be better addressed.

We have rechecked your comment, but we think that it is already clearly stated that the magnitude of a Dry-Hot events refers to the temperature while the duration of these refers to the dry spells:

"although in Europe the magnitude (temperature) of these events was revealed to have greater weight than their duration (dry spells) as indicated by Manning et al., (2019)."

Moving averages and future periods - I suggested the authors to explain the use of a 7-year moving average and the authors changed for 5-year moving year average without explaining. The authors could mention that in addition to the 5-year moving year average, they also used a 7-year moving average, and which were the main differences, if any.

There is no difference between 5-yr and 7-yr for the running window. We have used an odd number because there must be a central year; also 5-yr is half a decade.

Moreover, the authors have changed the analyzed future period from 2011-2100 to 2006-2100 without mentioned it in the review, and the three periods changed from (2011-2040, 2041-2070 and 2071-2100) to (2016-2035, 2046-2065 and 2081-2100). This should be stated clearer. The new future periods have now 19 years each (instead of 29 as before), and I think that the correct period to include in the abstract L13 is 2016-2100 instead of 2006-2100.

The authors would like to apologize for this oversight. The changes were marked in the text but not mentioned in the revision. The selection of these new periods is in accordance with the periods used in fifth IPCC (Stocker, 2015). The data from the RCPs starts in 2006. We used the complete series in Fig 10 and 11.