Response to RC1 (anonymous)

Thank you for your comments, we are glad you find the study interesting and appealing to a wide readership. We generally agree with your comments, including the main points that the motivation for the study should be better explained and that the discussion would benefit from significant shortening to focus on the core information. We are happy to implement the requested or suggested changes to the benefit of the paper.

Please find below our responses to the individual comments.

Comments

1. **Motivation.** The motivation for the study needs to be clearer. Currently the introduction reads like a series of definitions rather than a compelling story of why this study should be undertaken. There are two core areas where I think this could be improved:
    1. There have been quite a few studies that have used or compared drought indices (many are cited in your discussion) so what does this study offer that is novel and different?
    2. Why focus on Great Britain (which often isn't thought of as a country that experiences lots of droughts!) and how are your research questions relevant to this region?

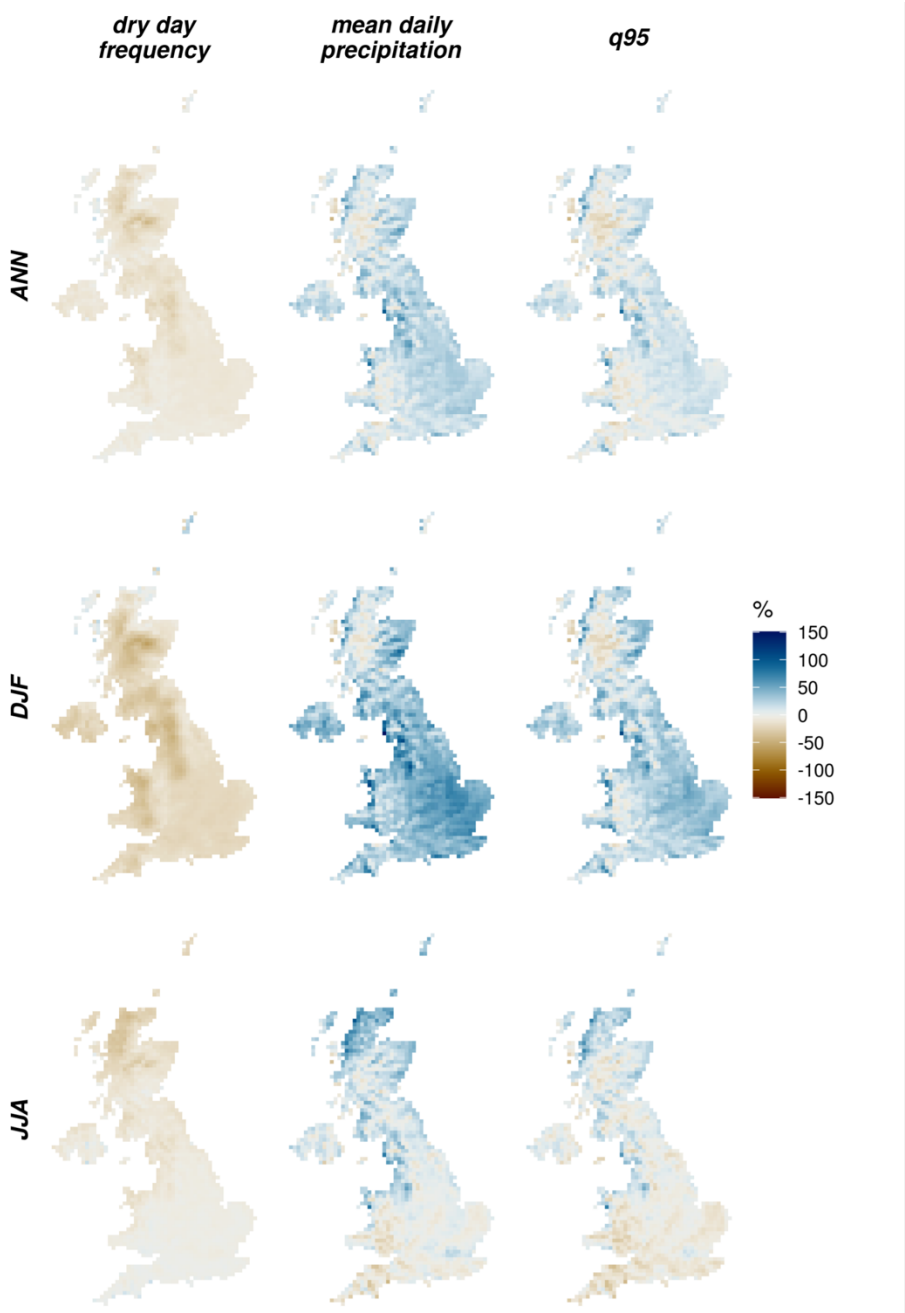
    We will take care to expand on these two core areas when writing a more compelling introduction. On 1.: indeed, many studies have compared drought indices, but we go in more depth here than is usually done, considering the differences in a wider range of drought characteristics. On 2.: First, the hydroclimatology of Great Britain is generally humid but still quite diverse, and drought is already a concern – especially in the South and East – as has also been shown in recent years. Moreover, water managers across Great Britain are facing challenges from different angles to secure adequate water supply in the future, including increases in demand and indeed climate change impacts. As this is a humid region, the primary concern with regards to future drought resilience is typically precipitation. This paper invites a critical examination of the potential importance of evaporation in a drier Great Britain due to climatic changes.
    Moreover, it points out just how different the resulting changes in a range of statistics are based on which atmospheric-based drought indicator was chosen to quantify droughts – arguably not surprising, but nevertheless a finding that is often underappreciated.

2. **Use of UKCP18.** There needs to be better justification for the use of the regional projections from UKCP18 in this study – why use this set of projections instead of the local UKCP18 projections for example? Or use a set of climate projections that encompasses different GCMs or RCP scenarios (for example)?
    We will improve the justification for the choice to use the UKC18 regional projections in our revised manuscript, along the following lines. The UKCP18 simulations are the de facto national projections for the UK, and have been produced with the aim of providing a spread of projections to support adaptation efforts in the UK. The local projections would have provided a greater added value if intense (convective) precipitation events on subdaily time scales or the added spatial resolution would have been crucial, however as droughts tend to be more spread out in space and time, we judged that the 12km daily resolution of the UKCP18 RCM pose a better trade-off between practicality and spatiotemporal detail for this purpose. While we agree that an ensemble that encompasses multiple GCM-RCM structures would add

another interesting dimension to the study, expanding the ensemble with e.g. EURO-CORDEX simulations would have been outside the scope and capacity of the study. The UCKP18 regional ensemble already samples parameter uncertainty in both the GCM and RCM, while our use of specific warming levels avoids a focus on particular RCP scenarios.

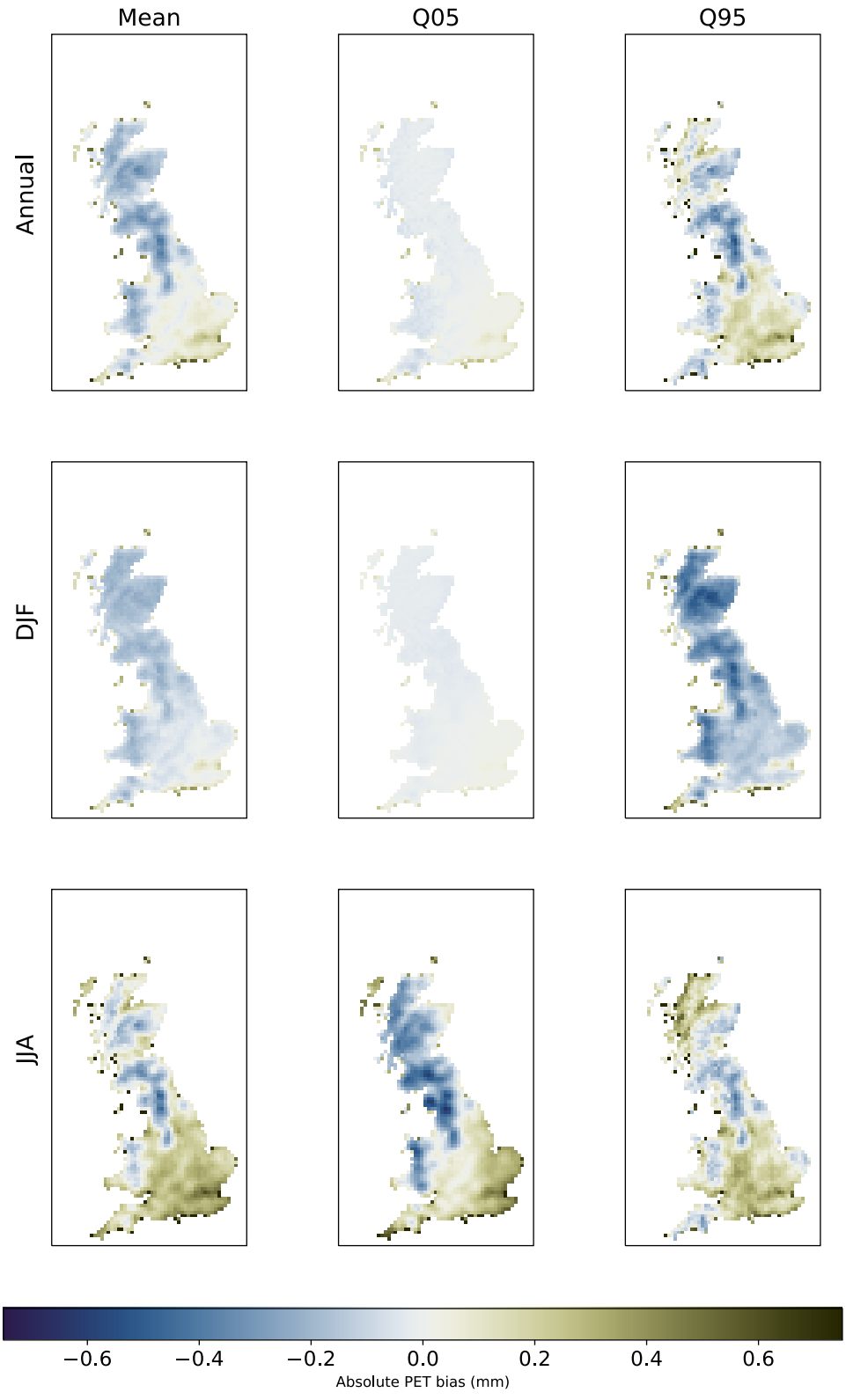
3. **Bias Correction.** At the end of Section 2.2 there is a section on bias correction which needs more detail. Why did you choose these bias correction methods and how are they applicable to the types of biases you observe between the climate projections and observational data? It would be helpful to add some plots in the supplementary information showing these biases to help the reader understand the nature of the biases and how they were corrected. For example, you note in section 3.2 that a direct comparison of the results between climate model ensemble members and observations is only possible because their distributions are similar – it would be helpful to see evidence of this.

   The focus of this study is not on the bias correction, though we agree that the reader would benefit from information about the model biases. The raw and bias corrected data were evaluated in detail. The results of this evaluation for temperature and precipitation will be presented in a separate paper about bias correction that is currently in preparation (for which the bias adjusted precipitation data will also be made available). We will include in the supplementary information of the current paper plots of the PET and precipitation biases (unless the separate bias correction paper is already available, in which case we will refer the reader to that paper for details of the precipitation biases). Nicole Forstenhäusler has kindly agreed we can include in this response (and the supplementary material of a potential revised version) some of the maps she produced for the bias correction paper.



Mean precipitation biases in UKCP18-RCM for 1981-2010, expressed as a percentage of the observed values. The bias for each ensemble member was computed and the mean across the ensemble is shown here. Dry-day frequency is the percentage of days with P < 1 mm; q95 is the 0.95 quantile of precipitation. Created by Nicole Forstenhäusler.

The maps we propose to include for PET will follow a similar approach, showing ensemble averaged biases in the daily PET mean, 5th and 95th percentiles in the Supplementary.



Mean PET biases (mm) in UKCP18-RCM for 1981-2010. The bias for each ensemble member was computed and the mean across the ensemble is shown here. Q05 and Q95 are the 0.05 quantile and the 0.95 quantile across.

The biases we observed for different quantiles were not equal to the biases observed in the mean, so we opted for a bias adjustment method that took this into account. Similarly, biases also varied between months and locations, so the bias adjustment needed to be specific for each month and grid cell. We considered more complex

methods that e.g. take into account different time scales or multivariate distributions, however it was unclear whether these methods would be beneficial compared to the univariate quantile mapping approach that is well-established in the literature. The ISIMIP3b method that was chosen in the end is based on quantile mapping, but also preserves projected changes in the variables being corrected, and adjusts the frequency of dry days separately – a desirable feature for drought research. In addition to the PET (and precipitation) evaluation plots in the Supplementary materials and the paper which will discuss the evaluation of raw and bias corrected precipitation (and temperature), we will also add the observation data to the following figures, so that the reader can see how they compare to the bias adjusted reference period UKCP-RCM simulations: Fig. 2 (aridity), Fig. 3 (average seasonal cycle), Fig. 4 (dots representing the observations on the scatter plots showing SPI6 vs SPEI6 drought frequency).

4. **Discussion.** The discussion section is extremely long and as a result, a lot of your interesting results get a little lost amid all the discussion. The authors need to have a critical read of the discussion and carefully consider if all the text is needed. As a suggestion, I would significantly shorten section 6.4 as this tends to be a review of the literature, rather than a discussion of your results with context from the literature (you could just use the summary paragraph – you don't really need all the preceding text).
   We agree that not all the text is needed to support the results and will significantly condense the discussion as such. Thank you for providing an example as a starting point.

Minor/Technical Corrections

P2 L45. 'Drought indices that only rely on atmospheric data are a popular choice due to data availability and **propagating model uncertainties**.' I don't understand this sentence – why are they a popular choice due to propagating model uncertainties?
Right, this sentence is skipping a few steps. We meant here that when precipitation and potential evapotranspiration are used to drive hydrological models, we need to make other choices that are uncertain (what (type of) hydrological model, how to calibrate it, …), and so atmospheric-based indicators provide a practical alternative. We will clarify this by replacing the last part of that sentence with "due to their ease of use (they do not require the deployment of an impact model, such as a hydrological model).".

P3 L65. You could add into the third research question the uncertainty in the RCM as a lot of your results focus on the differences between ensemble members. E.g. How sensitive are the projected changes in drought characteristics to the choice of atmosphere-based drought indicator and parametric uncertainty in regional climate models?
Thank you for pointing this out, we will include the element of comparing the sensitivity to the drought indicator choice with the sensitivity to the sampled RCM parameter uncertainty and the GMWL, as this is indeed something we put focus on.

P3 Section 2.1. It may be useful to state the time-period you used from each observational dataset.
We will add this in. From both datasets, we use 1961-2010 for the SI calibration, 1981-2010 for the bias correction and 1981-2005 for comparisons to the reference period in the results.

P5 L141-142. Why did you include aridity as well? What does it add to the story? The motivation could be a little clearer.
It accompanies the seasonal cycle plots of precipitation and potential evaporation in establishing an understanding of the mean climatic changes projected in UKCP18-RCM, before we discuss the drought frequency changes in depth, and providing a metric that is

more intuitive to interpret because it is based on physical rather than standardised quantities.

P6 L168. The area for London is small, but it is a central hub which contains around 14% of the population of GB and is likely to be significantly affected by droughts in the future. Consequently, leaving out the figures for London because the area is small seems an odd choice, given the significant impacts changes in droughts will have in this small part of GB. Potentially a better reason would be because the results are very similar the South-East region or East of England?
In response to a reviewer request to reduce the overall content of the paper, we will reduce some figures in the main text to show only 4 selected regions, with results showing all regions still included in the Supplementary information. We will then take this opportunity to reconsider inclusion of results for London in the latter, if they add relevant different information.

Figure 1. It is difficult to see the labels for North and West Scotland – can these be moved or made clearer?
We will adjust this.

P9 L203-205. This sentence is a little difficult to understand – can it be rewritten?
Agreed, we will adjust this. Proposed rephrasing:
"Using SPI6, drought frequency is projected to increase slightly on average under +2 °C above pre-industrial levels, with larger relative increases for the extremely dry conditions."

P9 L209. 'For extreme meteorological drought, all ensemble members **project multiples** of the reference period frequency by +4 ◦C.' I don't understand what you mean here.
Proposed rephrasing: For extreme meteorological drought, the projected frequency increases between two- and eightfold by +4°C compared to the reference period frequency, across the ensemble.

P10 L233 'due to SPEI6 occurrences beginning to saturate when they have already become quite frequent'. What do you mean by 'beginning to saturate'?
We will rephrase this. The point of this sentence was that, for the SPEI6 and the highest warming level, at some point the (vast) majority of summers are classified as drought, and so summer droughts cannot become much more frequent. As the summer droughts contribute the most to projected increases in drought frequency, the main contribution to the projected drought frequency changes starts to become "saturated" for SPEI.

P28 L666. I think you are missing some key references from the brackets?
Indeed, thank you for spotting, these will be added in.

Code and Data Availability – Great to see that the data you produced are available but this section needs to be a full description of all the data used in the study, including links to all the data you used for analysis (i.e. for the CHESS-PE, HadUK and UKCP18 data). I would also reword to 'The SPEI and SPI data **produced in this study** are available on Zenodo (doi:10.5281/zenodo.6123020) (Reyniers et al., 2022b) alongside the bias adjusted UKCP18- based PET (doi:10.5281/zenodo.6320707) (Reyniers et al., 2022a).'
Thank you very much for pointing this out, we will add this and reword accordingly.