Anonymous Referee 1
Author response

General comments
The submitted manuscript addresses a very relevant topic for water risk management, (i.e. low likelihood/high impact events) and does so using storylines, a novel approach that allows the investigation of plausible but unrealized high impact events. The selected storylines are based on the 2010-2012 UK drought event and explore imposed changes to 1) Precondition severity, 2) Temporal drought sequence, and 3) Climate change. The implications of such changes are assessed by quantifying changes to streamflow maximum intensity, mean deficit, and duration. The results do not only facilitate the realization that it could have been worse/it possibly will be worse but also sheds light on physical catchment properties that play a key role in the propagation of a multi-year drought event. In general, the manuscript is well written and structured and the results are relevant to a broad community interested in novel approaches that tackle environmental risk management and future climate change impacts. I have few minor concerns that I share in what follows:

RESPONSE: We thank the reviewer for the positive feedback on our manuscript. We are grateful for the comments and suggestions on how our manuscript can be improved. We respond to each comment given in the text below (in bold and italics text).

Specific comments
I understand plausibility to be a key property of the designed storylines. The first storyline proposes varying 3- and 6- months prior precipitation conditions to the 2010-2012 drought event independently of other climatic variables used in the model simulation. Such manipulations do not consider correlation structures in the data. I find that not completely justified and slightly weakening the plausibility assumption. For example, the potential presence of autocorrelation among successive monthly precipitation values or the correlation between precipitation and temperature are not considered. The authors can potentially mention these concerns in their discussion to further strengthen the plausibility argument.

RESPONSE: We agree that further information is needed to discuss the implications of the precondition storylines on the correlation between potential evapotranspiration (PET) and precipitation. Figure R1 (below) shows monthly precipitation and PET from 1965-2015. Apart from a slight negative correlation between precipitation and PET in spring and summer, there appears to be no clear correlation between the variables in the remaining months from 1965-2015 data. Figure R2 shows the equivalent values after precipitation 3- (i.e. OND 2009) and 6-months (i.e. JASOND 2009) before the 2010-12 drought reduced to match mean OND or JASOND precipitation at four return periods. The reduction to precipitation prior to the drought does not appear to be outliers compared to the observed relationship between precipitation and PET from 1965-2015. We also emphasize that the creation of event-based storylines in other locations should consider potential correlation between the different variables if a strong correlation is found. We will amend the text to reflect this. Figure R1 will be added in the supplementary materials and Figure R2 in the main text of the revised manuscript.

Regarding the reviewer's other concern, monthly precipitation values show very low autocorrelation among successive monthly precipitation values in all years (1965-2015). The average autocorrelation for successive monthly precipitation values across all years is -0.046 with values falling within the 95% confidence interval, indicating no statistical significance. As there is low autocorrelation, we believe the reductions to precipitation applied in the storylines of precondition severity are justified.

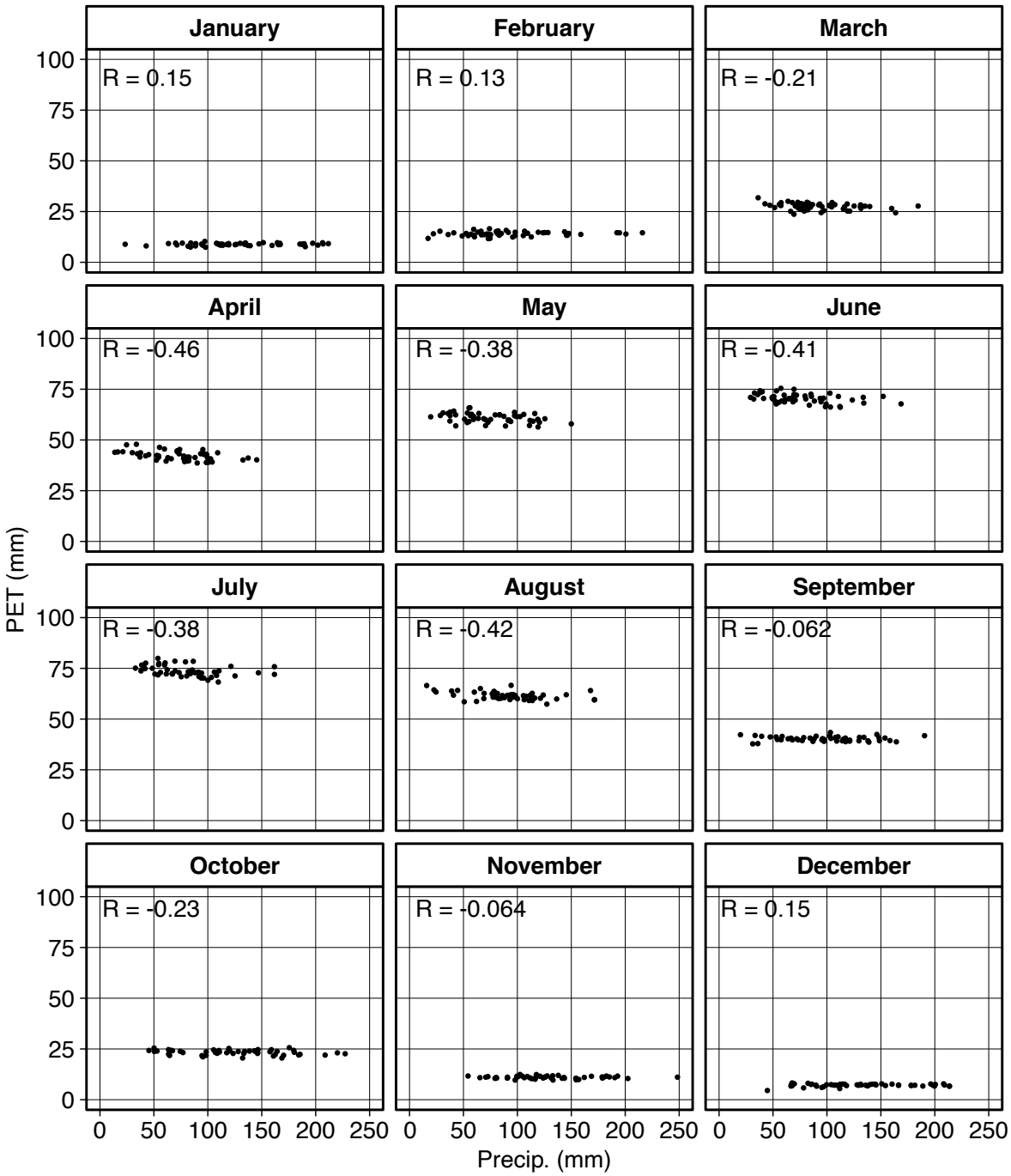

Figure R1 Observed relationship between PET and precipitation for each month for the period 1965-2015 averaged across the 100 UK catchments selected with the correlation coefficient value shown for each month.

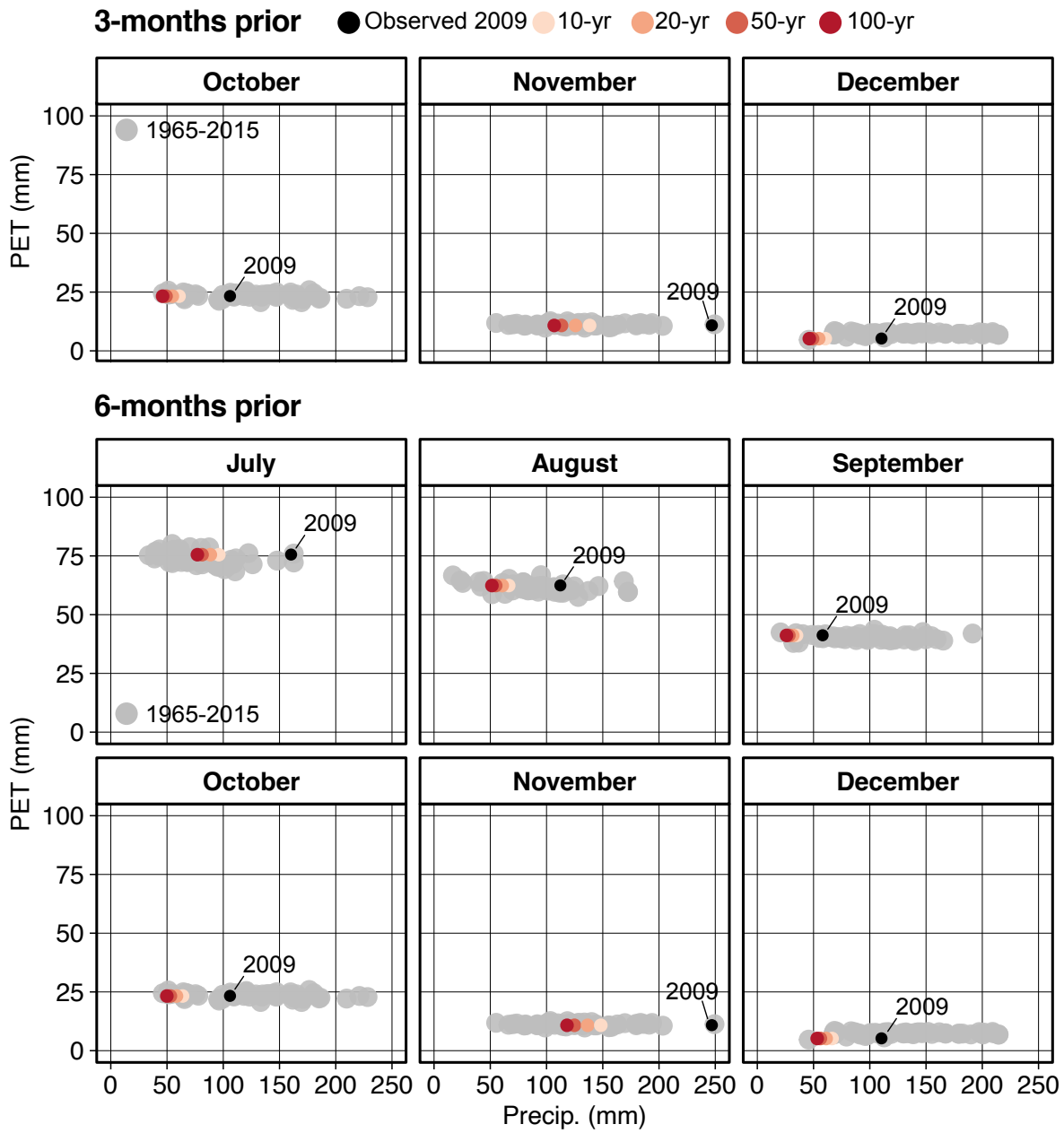

Figure R2 October to December monthly precipitation and PET (1965-2015) (top) and July to December monthly precipitation and PET (1965-2015) (bottom) The black circle indicates observed value in 2009 while the colored circles indicate the value after the precipitation 3- (top) and 6-months (bottom) prior to the 2010-12 drought is reduced at four return periods.

I see that some consideration is given in the paragraph starting at Line 516, nevertheless, I find that rather short and in itself not fully convincing. If I understand correctly, the authors address plausibility for the precondition storylines by comparing the resultant 12-month precipitation deficits to outputs of high-end climate change scenarios. They argue that the preconditioning storylines are plausible as these are contained within the range of outputs from high-end climate change scenarios. Nevertheless, I expected that plausibility concerning these particular storylines should address whether such conditions are possible in the current climate.

RESPONSE: The Environment Agency vulnerability framework and the high-end H++ climate change scenarios were intended as a point of comparison when

discussing the implications of the storylines of precondition severity instead of a justification of their plausibility. We will amend and move this text to Section 4.2 in our discussion of the value of the storyline approach to highlight these storylines as alternatives to existing projections. As discussed in the previous response, we have expanded our justification of the plausibility of the precondition storylines.

The authors state that they apply the delta approach in its standard form (line 189) where historical variability is retained. This formulation confuses me a bit as I am not sure what a non-standard form for the delta approach is.

RESPONSE: The standard form of the delta/change factor approach, as applied in this study, retains historical variability with monthly change factors. There have been different modifications or variations to the delta approach proposed in the literature. They mostly consist of ways to calculate percentile- or quantile-based change factors for relative changes in wet and dry days and rainfall intensity (e.g. Anandhi et al. 2011; Willems and Vrac 2011; Ntegeka et al. 2014). Anandhi et al. (2011) also reviews and presents a classification of different variants of the change factor method. Although there are several modifications, the standard delta method as used in this paper remains the most widely used. We will clarify this by referencing these studies in the revised manuscript. In response to the next comment, we will also expand on the limitations of this method and discuss other alternative methods.

Can the authors expand on this in their discussion to address limitations associated with the method they chose and possibly elaborate on other potential methods that can be used to answer questions such as: How would that particular event look like in a warmer world? (e.g. Wehrli et al. 2020).

RESPONSE: We will expand on limitations relating to the use of the delta method. The obvious limitation is that the delta method omits the influence of changes in wet/dry sequences. While this is a limitation to the storylines of climate change, the other storylines created in this study considers alternative changes to the wet/dry sequence of the observed drought.

Thanks for pointing us to Wehrli et al. (2020). We will add the citation as suggested as well as van Garderen et al. (2021) which is also an example of the use of atmospheric nudging as an alternative method to construct event-based storylines of extreme events in a warmer climate. Alternative approaches to investigate extreme events in a warmer world would be to search for analogues or events similar to the 2010-12 drought (for example, analysis of weather types or circulation patterns – e.g. Cattiaux et al. 2010, or through the use of large ensemble climate model data – e.g. van der Wiel et al. 2020). We will amend the text in the revised manuscript in the limitations and future work section to highlight these alternative approaches.

In response to a previous point, we will also add additional justification of the change factor method in the methods section. This particularly relates to challenges associated with the statistical bias correction and downscaling techniques, all of which assume that the biases corrected for and the bias

adjustment technique remain valid for future time periods. We believe the delta change method is suitable for this study as retaining the observed temporal sequence of the 2010-12 drought increases realism and enables quick comparison with the other storylines which were also created based on altering the observed time series of the 2010-12 drought.

It is clear to me why storylines are relevant as complementary information to already existing approaches that rely on GCM projections to quantify the hydrological impacts of climate change. I also do understand how these two approaches are very much different in scope. Nevertheless, the authors use the terms “scenario-driven approach” as a particular feature of GCM driven assessments in an attempt to contrast their approach and I find that slightly misleading. Storylines are still very much scenarios to my understanding, event-based in that case, and with a focus on plausibility rather than probability. I don’t see why they wouldn’t qualify as scenario-driven. The author themselves state that (i.e. line 143): “storylines follow similar methodologies employed in previous studies to create scenarios”. I, therefore, recommend revisiting specifically this phrasing to reduce confusion and facilitate the understanding of what is meant by storylines.

RESPONSE: The term “storylines” has been used in various ways in different disciplines. We will clarify our definition of “storylines” used in this paper. “Storylines” as used here refer to event-based “physical climate storylines” with a focus on plausible changes to causal elements of the 2010-12 drought. This differs from traditional GCM-driven climate change impact assessments which are constrained by GCM projections following different emissions or socio-economic scenarios (e.g. SSPs, RCPs etc.) when used as inputs to impact models. Their results are therefore constrained by the choice of the different scenarios and GCMs. “Physical climate storylines” can be created independently from these to represent situations or conditions that could lead to significant impacts and complement results from GCM-driven approaches. We will clarify the differences between existing approaches and our creation of “physical climate storylines” throughout the paper. We have also explained what was meant by our use of “scenario-driven” more explicitly.

We agree that line 143 may be confusing to readers. We consider the similar methodologies in the previous studies as cited to also be storylines that could be used to complement climate change projections and stress test hydrological systems. We will remove any mention of scenarios in this case.

Another point related to terminology: Can the authors explain their use of the term “counterfactual” when discussing future impacts of climate change. As the climate change storyline refers to a hypothetical event in the future, I find it a bit unclear why that would qualify as a counterfactual.

RESPONSE: We agree with the reviewer on the potentially confusing use of the term “counterfactual” when discussing climate change. We will remove any use of the term when discussing the use of the UKCP18 climate projections.

Technical comments

1. I am slightly confused by this sentence: Line 373, “The drought is estimated to worsen for the “Dry year before” storyline for all clusters except for mean drought deficit for Cluster 4 for SSI-6”. I believe something along the lines of: “The drought defined by SSI-6 is estimated to worsen for the “Dry year before” storyline for all clusters except for mean drought deficit for Cluster 4 ” is a bit more clear.
   RESPONSE: Thanks for the suggestion. We will modify the text as suggested in the revised paper.

2. Line 378: is -> are
   RESPONSE: We will change this in the revised paper.

3. Line 381-382: I believe something in the punctuation of the phrase is incorrect. Please check that.
   RESPONSE: We will modify this in the revised paper.

References

Anandhi, A., Frei, A., Pierson, D. C., Schneiderman, E. M., Zion, M. S., Lounsbury, D., and Matonse, A. H.: Examination of change factor methodologies for climate change impact assessment, Water Resources Research., 47, https://doi.org/10.1029/2010WR009104, 2011.

Ntegeka, V., Baguis, P., Roulin, E., and Willems, P.: Developing tailored climate change scenarios for hydrological impact assessments, Journal of Hydrology., 508, 307–321, https://doi.org/10.1016/j.jhydrol.2013.11.001, 2014.

van der Wiel, K., Selten, F. M., Bintanja, R., Blackport, R. and Screen, J. A.: Ensemble climate-impact modelling: extreme impacts from moderate meteorological conditions, Environ. Res. Lett., 15(3), 034050, doi:10.1088/1748-9326/ab7668, 2020.

van Garderen, L., Feser, F., and Shepherd, T. G.: A methodology for attributing the role of climate change in extreme events: a global spectrally nudged storyline, Natural Hazards and Earth System Sciences., 21, 171–186, https://doi.org/10.5194/nhess-21-171-2021, 2021.

Wehrli, K., Hauser, M., and Seneviratne, S. I.: Storylines of the 2018 Northern Hemisphere heatwave at pre-industrial and higher global warming levels, Earth Syst. Dynam., 11, 855–873, https://doi.org/10.5194/esd-11-855-2020, 2020.

Willems, P. and Vrac, M.: Statistical precipitation downscaling for small-scale hydrological impact investigations of climate change, Journal of Hydrology.,402, 193–205 https://doi.org/10.1016/j.jhydrol.2011.02.030, 2011.