Editor Comments

Thank you for responding to the three referees. They made several suggestions on how to improve the manuscript. In addition, the referees criticized the novelty of the paper and the fact that only one climate model projection (GCM) was used in the study. After reading your responses, I understand you do not intend to present "another case study of climate change impacts in the UK", but rather you wish to discuss the scientific implications of simulating hydrology under climate change using a gridded approach. The paper should be revised in light of the reviewers' comments and with any modifications you deem necessary to highlight the "scientific novelty" of your study. Together with the revised manuscript, please provide point-by-point responses to all reviewers' comments. I look forward to receiving the revised text.

Response: We have now revised the paper in light of these comments, please see a list of changes and our responses to each reviewer comment below. Comments are given in blue, and our responses given in black.

List of Changes

Line numbers refer to lines in the tracked changes version of the manuscript.

Section 2.3 (Lines 145-154), added further justification for using the UKCP18 climate projections.

Section 2.3 (Lines 156-160), added justification for use of the RCP8.5 emissions scenario.

Section 2.3 (Lines 270-220), added sentences on persistence bias, "Whilst potentially another interesting avenue of research in bias correction, namely wet/dry persistence bias, we decided not to pursue this analyses. Because we feel the matter is complex and requires a more dedicated paper on these issues and potential impacts, for example Moon et al. (2019) showed more wet/dry persistence biases between observed gridded rainfall products than between those and climate model outputs."

Section 2.4 (Lines 234-236), added sentence on missing snow processes in the model, "This model structure does not include a snow module, as snow processes were assumed not to substantially impact many GB catchments (95% of the catchments included in this study have less than 6% of precipitation falling as snow)."

Section 2.4 (Lines 253-255), explained selection of baseline and future periods, "These 25-year baseline and future periods were selected to allow the maximum distance between the baseline and future. The choice to start the baseline period in 1985 was due to the need for a long hydrological model spin-up period (1981-1985), which is required for some catchments in the south-east of England."

Section 2.5 (Line 267) amended "average flows" to "median flows."

Section 3.2 (Lines 296-299), clarified sentence on removal of reservoirs for the performance evaluation. Also referred to the new supplementary Section, "However, the presence of reservoirs was not found to lead to a reduction in model performance (see Supplement S2)."

Section 3.3 (Lines 326, 330,336), updated references to the supplement.

Section 3.4 (Lines 357-358) clarified sentence to refer to "median and high" flow changes only.

Section 3.5 (Lines 366-368), added reference to additional plots now given in the supplement.

Section 3.5 (Line 373) corrected spelling mistake (anticedent to antecedent).

Section 4.3 (Lines 456 – 458), changed sentence wording following reviewer advice, from saying metric selection was a source of uncertainty to highlighting the importance of considering multiple metrics.

Section 4.3 (Lines 464 – 467), added additional reference suggested by reviewer, "For example, Chegwidden et al., (2019) used an ensemble of two RCPs, 10 GCMs, two downscaling methods and four hydrological model structures in their analysis of climate change impacts on annual streamflow across the Pacific Northwest of North America, finding that GCMs were overall the dominant contributor to the variance in projected changes."

Section 4.3 (Line 479), added "future emissions scenario" into the list of uncertainty sources.

Section 4.3 (Line 483) recognised the limitation of using a single GCM, saying that "Other GCMs may have resulted in different precipitation trends into the future."

Supplement section S2, added new section to the supplement on "The impact of reservoirs on model performance." The following text and a new figure were added,

> "In the main text we evaluated the performance of the RCM-hydrological modelling chain. Catchments where runoff was affected by reservoirs or heavily regulated flows were excluded from this analysis, as the model does not simulate these processes and so errors in these catchments would likely not be due to the RCM data. Catchments impacted by reservoirs/regulated flows were identified using the factors affecting runoff (FAR) from the UK Hydrometric Register (Centre for Ecology and Hydrology, 2008). Here, we explore whether the presence of reservoirs/ heavily regulated flows has a large impact on model performance.
>
> Of the 346 catchments included in this study, the majority (60%) have no reservoirs in the catchment. 71 gauges (20%) have 1-5 reservoirs upstream, and 20% of gauges have more than 5 reservoirs upstream. While 40% of gauges do have a reservoir in the catchment, the capacity of the reservoirs is an additional important indicator of its impact on the flow time series as many of these reservoirs have a small capacity relative to the average precipitation and flow at the gauge. Of the 346 gauges, only 20 (5%) had a capacity greater than 10% of mean annual rainfall.
>
> Figure S13 shows the percentage error in RCM-driven simulations over the observed period. Catchments have been split into those with reservoirs/regulated flows (red) and those without (blue). When looking at median flows (Q50) there is no discernible difference in performance between catchments with and without reservoirs/ regulated flows. Surprisingly, when looking at high flows (Q10, Q1 and AMAX) the catchments with reservoirs or regulated flow regimes tend to have lower percentage errors. This could be due to their location in the country, with reservoirs often in wetter areas and therefore smaller percentage errors. Overall, these plots show that including catchments with reservoirs/ regulated flows would not have reduced the model performance presented in the main paper."

Supplement section S4, added new section to the supplement on "Additional plots showing the relationship between changes in precipitation and changes in flow." The following text, and two new figures, were added: "In the main paper we explored the relationship between 95th percentile precipitation change and Q1 change across all catchments, showing only the median of all hydrological parameter sets. Here we present additional plots looking at the relationship between changes in precipitation and changes in river flows for other precipitation quantiles (Figure S20), flow quantiles (Figure S20) and hydrological model parameter sets (Figure S21)."

Reviewer 1 Comments

RC: Lane et al. present an analysis of high flow metrics for 346 catchments in a future climate, considering 30 different parameter fields and 12 different RCMs. Although there is really nothing substantially wrong with the approach, especially when the goal is to support policy making, I think scientifically there are some missed opportunities which makes that the scientific gain of this study is currently limited.

Response: Thank you for your helpful feedback and for taking the time to write a review. Whilst we agree the results of the paper will be useful to support policy making, it is primarily a science based paper. As such, the paper explores the complex issues of uncertainty evaluation within national catchment responses under climate change. It highlights the value in considering hydrological model parameter uncertainties (found to be especially important for catchments in southeast England), and investigates how predicted responses to climatic change differ over regions and national scales that also depend on catchment characteristics and geoclimatic regimes.

RC: Several choices are not properly substantiated...

Response:

We do not agree with this assertion. Ultimately all climate impact modelling papers will have made different methodological choices and simplifications. Here, we focused on analysing uncertain climate impacts at the national scale, which is already computationally demanding within a high resolution (1.2km2 average HRU scale) distributed model. Additionally we evaluate the model parameter uncertainties and impacts of these cascades on streamflow changes. This goes beyond previous national studies for Britain, which have not previously considered hydrological parameter uncertainties over such a large number of catchments.

We respond to comments on each choice separately below. In response to these comments we have added further justifications of our choices in sections 2.3 and 2.4, as shown in the tracked changes manuscript.

RC: All RCMs are based on one single GCM. Why only one, and why this one? It reads as if this is not necessarily the GCM that performs best in the region of GB. Only RCP8.5 is considered, why only this one and what would that imply for the results?

Response:

The aim of this study was to explore hydrological model parameter uncertainties within a national climate impact study. We selected the UKCP18 climate projections to help us meet this aim as they have many advantages over other products, including 1) they were the nationally recognised highest resolution RCM climate model outputs available for a continuous time period over GB, 2) they were specifically developed for the UK and previous UKCP products have formed the basis of UK climate policy (Murphy et al. 2018), 3) they include a measure of climate uncertainty through the use of an RCM ensemble, 4) as RCM projections they are high resolution (12km) and have full spatial and temporal coherence which is needed to evaluate future climate change impacts on high flows in a spatially distributed hydrological model, 5) they are the newest national climate projections for GB, including the latest developments in climate modelling capability and scientific understanding, and have not yet been comprehensively analysed in other studies.

The UKCP18 projections only included RCM simulations for a single GCM, but still explored some climate uncertainties through the use of an RCM ensemble. This approach was also used for the UKCP09 climate projections which have been used in many UK climate impact studies (e.g. Prudhomme et al. 2013, Bell et al. 2016, Kay et al. 2018). The RCM ensemble was considered sufficient for our aim of assessing the hydrological model uncertainties within a national climate impact study. Importantly, we also found that minor differences between the RCM runs resulted in a huge variation of hydrological implications, showing that the RCM parameterisations which may be expected to be less influential were crucial after all. We are aware that the use of a different GCM would produce differing results, and this limitation is recognised in our discussion (lines 445-452). In response to all reviewers commenting on the use of a single GCM, we have explained this in section 2.3.

UKCP18 also only included RCM projections for the RCP8.5 scenario. We considered this the most important scenario to look at for two reasons: 1) it shows the 'worst case' and so will most likely show the largest expected changes, 2) the emissions in RCP8.5 are in close agreement with historical total cumulative CO_2 emissions and more and more are looking like a plausible future (Schwalm et al. 2020). But again we recognise that our results would have been different if an alternative scenario had been used, and we acknowledge that it is best to use multiple scenarios if the information is available.

In light of comments from all reviewers, we have expanded section 2.3 on Climate model data to highlight the above justifications, and have expanded on the discussion of missing uncertainty sources to include emissions scenario.

The following has been added on the choice of projections: "These projections were chosen because they have many advantages over other available products for UK impact assessments, including 1) they were the highest resolution (12km) RCM climate model outputs available for a continuous run period over GB, 2) they were specifically developed for the UK and form previous UKCP products have formed the basis of UK climate policy (Murphy et al., 2018), 3) they included a measure of climate uncertainty through the use of an RCM ensemble, 4) they are UK specific climate projection tools designed to help decision-makers assess their risk exposure to climate and thus compliment important evidence of climate impacts across the UK, 5) they were the newest national climate projections for GB, including the latest developments in climate modelling capability and scientific understanding, and therefore have not yet been comprehensively analysed in other impact studies. A key advantage of the RCM data over other UKCP18 products is that it has full spatial and temporal coherence and therefore allows for the assessment of interactions between changes in precipitation and PET as well as providing a nationally consistent picture of future changes (Met Office, 2020)."

The following has been added to explain why we used the RCP8.5 scenario: "The 12 RCM projections were all driven by the same GCM (GC3.05), and only the RCP8.5 emissions scenario was provided. We considered this to be the most important emissions scenario to look at for two reasons; 1) it shows the 'worst case' and so will most likely show the largest expected changes, and 2) the emissions in RCP8.5 are in close agreement with historical total cumulative CO2 emissions and are therefore increasingly looking like a plausible future (Schwalm et al., 2020)."

RC: DECIPHeR gives nice opportunities to test multiple model structures, why only one, and why only Topmodel?

Response:

We previously developed a modelling framework that uses spatial parameterisations which are nationally consistent and reflect core parametric uncertainties, constrained by available data using MPR (Lane et al. 2021). Our next step was to apply the model setup from Lane et al. (2021) to evaluate the hydrological modelling uncertainties when driving the model with a set of climate change scenarios. This manuscript presents the first GB-wide study to include hydrological model parameter uncertainties alongside RCM uncertainties, and further explores the relationship between catchment characteristics, climatic changes and changing high flows across a large sample of catchments. For this climate impact study, it was important to use the DECIPHeR model structure based on topographic flow gradients as a primary metric to define hydrological similarity as it has already been evaluated and parameterised across Great Britain (Coxon et al. 2019; Lane et al. 2021). We would like to note that DECIPHeR is not TOPMODEL except that topography is an important driver of hydrological response.

Whilst testing model structural variability was not an aim of this study, we agree with you that DECIPHeR gives a great opportunity to test different model structures and this is an ongoing research goal. Adding additional structures for different geologies, general climate variations, land management practices and human impacts is ongoing research that will take time to develop from our findings of Lane et al. (2021). However, that does not detract from the importance and need of quantifying uncertainties in hydrological impacts with our recent national model simulations.

RC: Why were different datasets used for P and PET for the bias correction? Both datasets seem to have both variables.

Response: The CHESS dataset uses CEH-GEAR as its rainfall data - we refer to the original dataset as is standard practice and to make clear the source of the data. The CEH-GEAR rainfall dataset does not include PET.

RC: Why was snow not included as process in the hydrological model? I was surprised that T was not required as input for the hydrological model (indeed, to simulate snow processes), to only find at the end of the discussion that snow was not accounted for - with a reference to eastern Scotland where the snow fraction can be up to 0.17. There might be valid reasons for each of these choices, but they can be better explained.

Response: Snow was not included as a process in the hydrological model because it affects so few catchments (95% of the catchments included in this study have less than 6% of precipitation falling as snow). The exception to this is in the Cairngorm mountains in Scotland where the fraction of precipitation falling as snow can reach 17% (Coxon et al. 2020).Consequently, we included this in the discussion so that readers were fully aware of this limitation.

We have added the following to the manuscript: “This model structure does not include a snow module, as snow processes were assumed not to substantially impact many GB catchments (95% of the catchments included in this study have less than 6% of precipitation falling as snow).”

RC: At several places, it is mentioned that caution should be taken when interpreting the results (related to snow, see later point, and to potential groundwater flow). Not even widely discussed are the catchments with reservoirs/regulated flow (they are only mentioned as being excluded for the analysis of the evaluation of the model chain, but how many of the 346 are (heavily) regulated? are they spatially clustered? and how can we somehow validate the simulation of these catchments if the regulation is not included in the model structure?). Taken into account these three factors (snow, groundwater, regulation) that require caution in the interpretation, it becomes a bit difficult to determine which numbers have meaning, and which don't. Given that uncertainty estimation is one aspect of this study, it are precisely these kind of aspects that might need more attention.

Response:

Of the 346 catchments analysed, the majority (60%) have no reservoirs in the catchment. 71 gauges (20%) have 1-5 reservoirs upstream, and 20% of gauges have more than 5 reservoirs upstream. While 40% of gauges do have a reservoir in the catchment, the capacity of the reservoirs is an additional important indicator of its impact on the flow time series as many of these reservoirs have a small capacity relative to the average precipitation and flow at the gauge. Of the 346 gauges, only 20 (5%) had a capacity greater than 10% of mean annual rainfall.

We excluded catchments with reservoirs/ regulated flows from the evaluation of the model chain because we would not expect the model to accurately simulate these flows without reservoir information in the model structure. However, in response to this comment we explored how the model error in simulating AMAX, Q1, Q10 and Q50 varied between the catchments with regulated flows/ reservoirs and those without. We found that there was not a reduction in performance between the catchments with regulated flows/reservoirs and those without.

This further analysis has now been added into the supplementary information – please see the new supplement section S2.

RC: Besides that many choices are not well substantiated, I think there are some missed opportunities in analyzing and presenting the effect of uncertainty (in this case, introduced by RCMs and parameters). For instance in Figure 7, one could expect that the parameters have influence on the non-linear relation between precip and Q1, while it is precisely the median of the parameter sets that is displayed here (same for the runoff coefficient). Therefore, after reading the manuscript, I still don't have the feeling I fully comprehend the uncertainty in the projections and their implications for the results.

Response: Again, we do not agree that ‘many’ of our choices are not well substantiated. When creating Figure 7 we explored using different hydrological parameter sets, and ended up selecting the median as the choice of parameter set made little difference to the overall picture. We have added plots to the supplementary information to demonstrate how this plot changes for different hydrological model parameter set selections, as well as to demonstrate how this plot differs for other precipitation quantiles and flow statistics. Please see the new supplementary section S4.

RC: In spite of what is written in the introduction ("Many studies have attempted to quantify the impact of these uncertainties by using multiple GCMs/RCMs, bias correction techniques, hydrological model structures and/or hydrological model parameter sets and propagating these uncertainties through the modelling chain. However, these studies are often focused on small catchment samples as the large numbers of simulations needed are computationally demanding"), there are studies that sample several steps of the modelling chain for a large sample of catchments (see references below, sorry to refer to my own work), and which might be useful in the discussion to put the results into perspective (how do these results compare to accounting for different GCMs and different hydrological model structures?). *Melsen et al. "Mapping disagreement in hydrologic projections", HESS 2018

*Chegwidden et al., "How Do Modeling Decisions Affect the Spread Among Hydrologic Climate Change Projections? Exploring a Large Ensemble of Simulations Across a Diversity of Hydroclimates", Earths Future 2019,

*Queen et al., "Ubiquitous increases in flood magnitude in the Columbia River basin under climate change", HESS, 2021

Response: Thank you for highlighting these papers. We note that two of these studies are focused on a single multi-nested river basin, albeit of considerable size. And that Melsen et al. focused purely on the sign of change and did not use any form of weightings to define the quality and fit of the ensembles of model simulations. We have added to the discussion section 4.3, "Previous studies found hydrological modelling uncertainties to be small relative to climate modelling uncertainties, especially when considering high flows (Chegwidden et al., 2019; Chen et al., 2011; Velázquez et al., 2013). For example, Chegwidden et al., (2019) used an ensemble of two RCPs, 10 GCMs, two downscaling methods and four hydrological model structures in their analysis of climate change impacts on annual streamflow across the Pacific Northwest of North America, finding that GCMs were overall the dominant contributor to the variance in projected changes."

RC: Last minor thing: In the discussion (l. 430) the selection of a metric is referred to as a source of uncertainty. I'm not sure I entirely agree with that. Different metrics will lead to different results, simply because they evaluate different things. That is, in my opinion, not uncertainty but simply the result of a (hopefully deliberate) choice. It does show, however, that it can be useful to evaluate multiple metrics.

Response: We agree with you on this. We have altered the text from "The selection of metrics used to explore climate impacts was a further source of uncertainty; the picture of climate change impacts on flows differed between the four metrics presented here." to "The overall picture of climate change impact on flows differed between the four selected metrics, showing the importance of metric selection and consideration of multiple metrics in model evaluation and impact studies."

RC: l. 244; average flow is not necessarily equal to median flow.

Response: We agree and have changed ‘average’ to ‘median.’

RC: If this review sounds a bit harsh, it is because I know the authors can do better. Most of the material is already there, therefore I am confident that the authors will be able to improve the manuscript such that it will add more to the scientific literature.

Response: We appreciate the thought and suggestions that have gone into your review. Of course we feel more strongly that our contribution highlights a number of core aspects and delves into the uncertainty problem nationally in a way that has rarely (and never in the case of UK climate simulations) been seen.

References:

Bell, V. A., Kay, A. L., Davies, H. N., & Jones, R. G. (2016). An assessment of the possible impacts of climate change on snow and peak river flows across Britain. Climatic Change, 136(3), 539-553.

Coxon, G., Freer, J., Lane, R., Dunne, T., Knoben, W. J., Howden, N. J., ... & Woods, R. (2019). DECIPHeR v1: dynamic fluxEs and connectivity for predictions of HydRology. Geoscientific Model Development, 12(6), 2285-2306.

Coxon, G., Addor, N., Bloomfield, J. P., Freer, J., Fry, M., Hannaford, J., ... & Woods, R. (2020). CAMELS-GB: Hydrometeorological time series and landscape attributes for 671 catchments in Great Britain. Earth System Science Data, 12(4), 2459-2483.

Kay, A. L., Bell, V. A., Guillod, B. P., Jones, R. G., & Rudd, A. C. (2018). National-scale analysis of low flow frequency: historical trends and potential future changes. Climatic Change, 147(3), 585-599.

Lane, R. A., Freer, J. E., Coxon, G., & Wagener, T. (2021). Incorporating Uncertainty into Multiscale Parameter Regionalisation to Evaluate the Performance of Nationally Consistent Parameter Fields for a Hydrological Model. Water Resources Research, e2020WR0283

Murphy, J. M., Harris, G. R., Sexton, D. M. H., Kendon, E. J., Bett, P. E., Clark, R. T., ... & Yamazaki, S. T. (2018). UKCP18 land projections: Science report.

Prudhomme, C., Haxton, T., Crooks, S., Jackson, C., Barkwith, A., Williamson, J., ... & Watts, G. (2013). Future Flows Hydrology: an ensemble of daily river flow and monthly groundwater levels for use for climate change impact assessment across Great Britain. Earth System Science Data, 5(1), 101-107.

Schwalm, C. R., Glendon, S., & Duffy, P. B. (2020). RCP8. 5 tracks cumulative CO2 emissions. Proceedings of the National Academy of Sciences, 117(33), 19656-19657.

Reviewer 2 Comments

RC: I reviewed the manuscript “A large-sample investigation into uncertain climate change impacts on high flows across Great Britain” by Lane et al. The main goal of this work is to present climate change impact projections on high flows for GB including climate model and hydrological model parameter uncertainties. There is nothing substantially new in this study: national scale studies already exist, studies with more options considered for each step of the modelling chain exist, and studies with a more thorough evaluation of the uncertainties exist. Overall, the study has some potential. Unfortunately, some methodological choices are questionable and do not reflect the state-of-the-art. Namely, the fact that one RCP only is used, one GCM only is used, that periods are not 30-year long and the reference period overlaps the "future" of RCPs is disturbing. Some figures are not properly introduced (I am not sure I understood well what is presented based on the captions). To summarize, I must say I failed to learn new insight neither on a methodological point of view nor on the future of GB floods (as one GCM and one RCP are used only). In addition, I second all comments from reviewer 1. Below are my major remarks, as well as miscellaneous minor or moderate points to tackle.

Response: Thank you for taking the time to review our paper and for your feedback, we have responded to all of the comments below. We disagree that we do not present anything substantially new. This study provides the first spatially consistent GB projections including both climate ensembles and hydrological model parameter uncertainties.

RC: A single GCM is used. I was very surprised to discover that the authors made the choice of using a single GCM. That's a rather unusual set up for this analysis: GCMs account for a large part of uncertainties in hydrological projections! The justification that comes at lines 145-147 (I understand it as the will to only sample the warmer range of possible climate outputs) is not convincing to me: another GCM, with a rather similar temperature pattern, might result in a very different precipitation evolution in the future - not all GCMs will result in intensified precipitation with similar spatial patterns, intensity and seasonality. As a consequence, the use of this single GCM might be interesting to estimate the warmer range of future temperature, but not at all to estimate the future range of hydrological variables!

Response:

We defer to the same responses we provided to Reviewer 1. In light of comments from all reviewers, we have expanded section 2.3 on Climate model data.

The following has been added on the choice of projections: "These projections were chosen because they have many advantages over other available products for UK impact assessments, including 1) they were the highest resolution (12km) RCM climate model outputs available for a continuous run period over GB, 2) they were specifically developed for the UK and form previous UKCP products have formed the basis of UK climate policy (Murphy et al., 2018), 3) they included a measure of climate uncertainty through the use of an RCM ensemble, 4) they are UK specific climate projection tools designed to help decision-makers assess their risk exposure to climate and thus compliment important evidence of climate impacts across the UK, 5) they were the newest national climate projections for GB, including the latest developments in climate modelling capability and scientific understanding, and therefore have not yet been comprehensively analysed in other impact studies. A key advantage of the RCM data over other UKCP18 products is that it has full spatial and temporal coherence and therefore allows for the assessment of interactions between changes in precipitation and PET as well as providing a nationally consistent picture of future changes (Met Office, 2020)."

The following has been added to explain why we used the RCP8.5 scenario: "The 12 RCM projections were all driven by the same GCM (GC3.05), and only the RCP8.5 emissions scenario was provided. We considered this to be the most important emissions scenario to look at for two reasons; 1) it shows the 'worst case' and so will most likely show the largest expected changes, and 2) the emissions in RCP8.5 are in close agreement with historical total cumulative CO_2 emissions and are therefore increasingly looking like a plausible future (Schwalm et al., 2020)."

RC: Choice of the study periods. The authors state that “changes in flow metrics between the baseline (1985 –2010) and future (2050 –2075) periods were evaluated.”. Why these choices? These are quite unusual for several reasons:

- the baseline contains 5 years with "future" GHG emission trajectories (2006-2010)

- the periods do not cover the classical 30-year period used to estimate climatology (WMO recommendation)

In addition, nowadays many studies assess the impact of climate change over the whole future period, using moving windows, which is a clear step forward as it allows to identify emergence times and trends.

Warmup period: Line 233: I find this justification rather disappointing. There are many ways to avoid throwing 5 years of climate data during the pre-RCP period (i.e. < 2005):

- using RCM data < 1981 for warmup (but I understand that it was not made available by Met Office, which is very surprising as most climate models usually start in 1950 or 1970 at least)

- using observed climate data from <1981 for warmup

- recycling RCM data from 1981-1985 for obtaining 1981 initial states

Response:

The 25-year baseline and future periods were selected to allow the maximum distance between the baseline and future.

The choice of a 1985-2010 baseline was due to 1) the need for a long model spin-up period as this is required for some catchments in the south-east of England, 2) the choice to have baseline and future periods that were at least 25 years long.

Thank you for your suggestions on how to reduce the need for a long model spin-up, we will consider these options for future work and of course would have used RCM data pre-1981 if it had been available. You are correct that the 2006-2010 climate data does include an element of predictive information, but 1981-2010 is the standard baseline of the WMO and is offered as an alternative baseline for users of the UKCP18 products (Murphy et al. 2018).

We agree that moving window analysis enables an interesting analysis of trends, and will consider this for future work. However, our aim was to present a national overview of potential changes using a spatial model, and this is easiest to visualise using baseline and future time-slices.

In response to this comment, we have added a justification of our choice to use these time periods into section 2.4 “These 25-year baseline and future periods were selected to allow the maximum distance between the baseline and future. The choice to start the baseline period in 1985 was due to the need for a long hydrological model spin-up period (1981-1985), which is required for some catchments in the south-east of England.”

RC: *Partially wrong assertions*: Line 331: This assertion is partially erroneous for me. The sources of uncertainties depend a lot on the indicator that is studied. This is not necessarily true for low flows, as shown in more recent studies as those cited by the authors.

Response: Thank you for highlighting this and for your suggested references, we agree with your comment. We have amended this sentence to "RCM parameters were a larger source of uncertainty in median and high flow changes than hydrological model parameters (see Figure 6). This finding agrees with previous studies that have investigated high flows, which generally find climate models to be the largest source of uncertainty in hydrological climate impact assessments (Addor et al., 2014; Bosshard et al., 2013; Kay et al., 2009)."

RC: In addition, RCPs also represent an important source of uncertainty that is not considered in this study.

Response: See our response to the 'A single GCM is used' comment above.

References:

Bell, V. A., Kay, A. L., Davies, H. N., & Jones, R. G. (2016). An assessment of the possible impacts of climate change on snow and peak river flows across Britain. Climatic Change, 136(3), 539-553.

Kay, A. L., Bell, V. A., Guillod, B. P., Jones, R. G., & Rudd, A. C. (2018). National-scale analysis of low flow frequency: historical trends and potential future changes. Climatic Change, 147(3), 585-599.

Murphy, J. M., Harris, G. R., Sexton, D. M. H., Kendon, E. J., Bett, P. E., Clark, R. T., ... & Yamazaki, S. T. (2018). UKCP18 land projections: Science report.

Prudhomme, C., Haxton, T., Crooks, S., Jackson, C., Barkwith, A., Williamson, J., ... & Watts, G. (2013). Future Flows Hydrology: an ensemble of daily river flow and monthly groundwater levels for use for climate change impact assessment across Great Britain. Earth System Science Data, 5(1), 101-107.

Schwalm, C. R., Glendon, S., & Duffy, P. B. (2020). RCP8. 5 tracks cumulative CO2 emissions. Proceedings of the National Academy of Sciences, 117(33), 19656-19657.

Reviewer 3 Comments

RC: This is a well written paper by a strong team of researchers.

Response: Thank you for taking the time to review our manuscript and for your suggestions.

RC: My main comment is on their Section 2.3 - The authors are limiting their investigation by using a single GCM and scenario for projecting into the future.

Response:

We defer to the same responses we provided to reviewer 1.In light of comments from all reviewers, we have expanded section 2.3 on Climate model data.

The following has been added on the choice of projections: "These projections were chosen because they have many advantages over other available products for UK impact assessments, including 1) they were the highest resolution (12km) RCM climate model outputs available for a continuous run period over GB, 2) they were specifically developed for the UK and form previous UKCP products have formed the basis of UK climate policy (Murphy et al., 2018), 3) they included a measure of climate uncertainty through the use of an RCM ensemble, 4) they are UK specific climate projection tools designed to help decision-makers assess their risk exposure to climate and thus compliment important evidence of climate impacts across the UK, 5) they were the newest national climate projections for GB, including the latest developments in climate modelling capability and scientific understanding, and therefore have not yet been comprehensively analysed in other impact studies. A key advantage of the RCM data over other UKCP18 products is that it has full spatial and temporal coherence and therefore allows for the assessment of interactions between changes in precipitation and PET as well as providing a nationally consistent picture of future changes (Met Office, 2020)."

The following has been added to explain why we used the RCP8.5 scenario: "The 12 RCM projections were all driven by the same GCM (GC3.05), and only the RCP8.5 emissions scenario was provided. We considered this to be the most important emissions scenario to look at for two reasons; 1) it shows the 'worst case' and so will most likely show the largest expected changes, and 2) the emissions in RCP8.5 are in close agreement with historical total cumulative CO_2 emissions and are therefore increasingly looking like a plausible future (Schwalm et al., 2020)."

RC: RCMs are essentially interpolators although with good physical realism. They work off lower and lateral boundaries that are simulated by the GCM. These lower and lateral boundaries have considerable biases unfortunately, and I often feel the use of RCMs without accounting for these biases as doing disservice to the use the simulation may have. We know that RCMs are well grounded to the topography, and hence one needs to question whether the RCM simulations with biased boundary conditions are essentially simulating the effect of the topography or the true change warming is about to unfold. There are two ways of correcting this limitation and unfortunately, both require a bit of work on the part of the authors. The first of these is to remove systematic biases in the lateral and lower boundaries that form the inputs into the RCMs. By this I am not referring to the post-processing bias correction the authors have perfomed here, but the bias in the boundary conditions before the RCM is run. This, however, requires new RCM runs which is a significant effort in terms of computing and time. The authors may want to go through the papers below that illustrate how useful this can be:

Rocheta, E., Evans, J. P. & Sharma, A. Correcting lateral boundary biases in regional climate modelling: the effect of the relaxation zone. Climate Dynamics 55, 2511-2521, doi:10.1007/s00382-020-05393-1 (2020).

Kim, Y., Evans, J. P., Sharma, A. & Rocheta, E., Geophysical Research Letters, Spatial, temporal, and multivariate bias in regional climate model simulations. 48, e2020GL092058 (2021).

The other way of addressing this limitation is to use multiple GCMs as the basis for boundary variables that feed into the RCM. While this does not address the biased boundary inputs the RCMs is subject to, it at least produces an envelope of the uncertainty that results from the use of a single (biased) GCM. This is the approach most researchers use in climate change assessment, often coupled with a post-processing step involving bias correction. I realise the authors may be limited in their access to RCM simulations from other GCMs, but, if so, need to at least discuss the implications this may have on their overall findings.

Response:

Thank you for these suggestions. We are not running the RCMs ourselves, so while this is an interesting idea it is not feasible in the context of our study. We do note in the discussion that a single GCM was used and therefore we do not sample the full range of climate uncertainty.

RC: The only other comment I have is regarding the use of the distributional bias correction adopted. Given the importance of antecedent conditions (which the authors have noted), not considering bias in persistence attributes can misrepresent the relationship between pre-storm wetness and storm extremes. This is evident even in urban catchments where one would usually not expect antecedent conditions to matter. It may be worthwhile for the authors to assess this dependence in their bias corrected precipitations and observations, incase there is a bias present. This may be especially important in those catchments where they are seeing a decrease in flood magnitudes.

Response:

We agree that bias in persistence attributes is an important and challenging issue. While bias in precipitation persistence has been evaluated in the climate modelling community (Armal et al. 2020; Kumar et al. 2013; Moon et al. 2019), we have not seen any papers where it has been addressed for hydrological modelling.

The choice of bias correction methodology is difficult, with a balance between correcting the data as much as necessary for meaningful hydrological simulations, but not making those methodologies too complex and thus even more difficult to justify as being plausible in future climates. We selected the distributional bias correction approach as it suitably corrects both high and low precipitation events (opposed to for example monthly mean bias correction which simply scales both light and heavy rainfall by the same percentage). This was successful in correcting the overall distribution of precipitation, including the mean daily rainfall bias and heavy rainfall biases discussed in the supplement S1. The bias correction we applied also significantly improved RCM overestimation of the number of rainy days (Figure S4), resulting in the RCMs having a similar number of wet days to the observations. It's important to recognise here that we cannot correct weather sequences that are likely to hold in future scenarios. We can only readily correct a statistical distribution approach to climate model output, indeed some colleagues do not correct at all and only look at the delta changes between current and future periods (e.g. Bell et al. 2009, Bell et al. 2012). Whilst we recognise the dangers in doing that (i.e. that you might be in a different non-linear rainfall-runoff response of catchment states by using uncorrected simulations), at least such approaches are not changing the physics within the models. Here we believe, for current best practice, we have taken the rational approach to statistical correcting that has improved wet/dry day RCM outputs and thus in part has improved such biases. We believe exploring these issues further would take a considerable amount of research to understand the nuances and impacts and how persistence bias changes are relevant for future scenario simulations, this is a very challenging piece of research. We will write more about this in our revised manuscript and discuss the issues. Finally, and we note this in our revisions of the manuscript, studies have shown that variations in persistence biases can be more variable between observed rainfall products than between these products and climate model outputs (Moon et al., 2019), thus again emphasising the difficulty of dealing with these types of biases and the best approach.

In response to this comment, we have noted that climate models may well have biases in persistence attributes, and therefore may not be able to simulate persistent wet/dry periods, but that rainfall observation products also have significant differences in wet/dry persistence, adding the following to section 2.3, "Whilst potentially another interesting avenue of research in bias correction, namely wet/dry persistence bias, we decided not to pursue this analyses. Because we feel the matter is complex and requires a more dedicated paper on these issues and potential impacts, for example Moon et al. (2019) showed more wet/dry persistence biases between observed gridded rainfall products than between those and climate model outputs."

References:

Armal, S., Devineni, N., Krakauer, N. Y., & Khanbilvardi, R. (2020). Simulating precipitation in the Northeast United States using a climate-informed K-nearest neighbour algorithm. *Hydrological Processes*, *34*(20), 3966-3980.

Bell, V. A., Kay, A. L., Jones, R. G., Moore, R. J., & Reynard, N. S. (2009). Use of soil data in a grid-based hydrological model to estimate spatial variation in changing flood risk across the UK. *Journal of Hydrology*, *377*(3-4), 335-350.

Bell, V. A., Kay, A. L., Cole, S. J., Jones, R. G., Moore, R. J., & Reynard, N. S. (2012). How might climate change affect river flows across the Thames Basin? An area-wide analysis using the UKCP09 Regional Climate Model ensemble. *Journal of Hydrology*, *442*, 89-104.

Bell, V. A., Kay, A. L., Davies, H. N., & Jones, R. G. (2016). An assessment of the possible impacts of climate change on snow and peak river flows across Britain. Climatic Change, 136(3), 539-553.

Kay, A. L., Bell, V. A., Guillod, B. P., Jones, R. G., & Rudd, A. C. (2018). National-scale analysis of low flow frequency: historical trends and potential future changes. Climatic Change, 147(3), 585-599.

Kumar, Sanjiv, Venkatesh Merwade, James L. Kinter III, and Dev Niyogi. "Evaluation of temperature and precipitation trends and long-term persistence in CMIP5 twentieth-century climate simulations." *Journal of Climate* 26, no. 12 (2013): 4168-4185.

Moon, Heewon, Lukas Gudmundsson, Benoit P. Guillod, Vuruputur Venugopal, and Sonia I. Seneviratne. "Intercomparison of daily precipitation persistence in multiple global observations and climate models." *Environmental Research Letters* 14, no. 10 (2019): 105009.

Murphy, J. M., Harris, G. R., Sexton, D. M. H., Kendon, E. J., Bett, P. E., Clark, R. T., ... & Yamazaki, S. T. (2018). UKCP18 land projections: Science report.

Prudhomme, C., Haxton, T., Crooks, S., Jackson, C., Barkwith, A., Williamson, J., ... & Watts, G. (2013). Future Flows Hydrology: an ensemble of daily river flow and monthly groundwater levels for use for climate change impact assessment across Great Britain. Earth System Science Data, 5(1), 101-107.

Schwalm, C. R., Glendon, S., & Duffy, P. B. (2020). RCP8. 5 tracks cumulative CO2 emissions. Proceedings of the National Academy of Sciences, 117(33), 19656-19657.