Reply to the Editor's comments

1. I don't think that is a relevant reference; CMIP5 is essentially an MME, not a PPE.

- (Reply) We have revised the sentence as follows.
- (Before) The first is multi-model ensembles (MMEs) method to address the structural uncertainty associated with the understanding and parameterisation of the GCMs. The second is the perturbed-physics ensembles (PPEs) method which is complementary to the MME approach, and is applied in the Intergovernmental Panel on Climate Change (IPCC) assessments (Meehl et al., 2007; Solomon, 2007; Taylor et al., 2012).
- (After) The first is a multi-model ensemble (MME) method to address the structural uncertainty associated with the understanding and parameterisation of the GCMs which is applied in the Intergovernmental Panel on Climate Change (IPCC) assessments (Meehl et al., 2007; Solomon, 2007; Taylor et al., 2012). The second is the perturbed-physics ensemble (PPE) method which is generated by perturbing physical parameters in a given climate model and is complementary to the MME approach.

2. It's the same issue, isn't it?

- (Reply) Yes, we have revised the sentence as follows.
- (Before) Another issue presented in this study is associated with how to correct the PPEs' bias to preserve the spread.
- (After) The primary research question presented in this study, hence, is associated with how to correct the PPEs' bias to preserve the spread.

3. Again, this is confusing. You claim that the observations are just one realisation of the real world, so they cannot be used to describe the real world. So your ensemble spread should NOT be similar to the observational spread.

- (Reply) Since the observations are just a possible realisation (i.e. sample space) of the real world (i.e. population), we have used a bootstrap method to make inferences about the natural variability from 30 year of observation data (a recognised method in computational statistics to estimate key statistical parameters). The basic concept of the proposed bias correction scheme is that the ensemble spread is maintained to a certain degree after bias correction which is compatible with the natural variability of the observation. However, it is important to alert to readers that this method has its limitation if the natural variability is not captured by the 30 years of observations (e.g., oscillations of very long return periods and unprecedented extremes). Such a limitation and possible solutions have been discussed in the revised manuscript in the Discussion section.

4. "The synthetic weather variables from WG have statistical properties ***similar to the observations*** since the WG is calibrated on the observations." => but not necessarily similar to the real world

- (Reply) We have revised the sentence as follows.
- (Before) The synthetic weather variables from WG have statistical properties similar to the observations since the WG is calibrated on the observations.

- (After) The synthetic weather variables from WG have statistical properties similar to the observations (although not necessarily similar to the real world) since the WG is calibrated on the observations.

5. "In this study we have focused on the mean value and the spread of bias-corrected RCM precipitation." Why? I think you need to make a reference to your hydrological case study here.
- (Reply) We have revised the text as follows.
  In this study we have focused on the use of the spread of bias-corrected RCM precipitation to investigate the impact of the conventional and proposed bias correction schemes on the flow. The conventional method removes the spread of the ensemble while the proposed method can better convey the spread properties of the ensemble.

6. I can see a rotation but not a rotation of 0.22 degree. Rotation relative to what?
- (Reply) There was a mistake in the manuscript. The 0.22° refers to the resolution of the grid. We have corrected it as follows.
- (Before) The temporal and spatial resolutions of the HadRM3 climate data are daily and 25km respectively. As presented in Figure 1, the RCM grid boxes are rotated by 0.22°.
- (After) The temporal and spatial resolutions of the HadRM3 climate data are daily and 25km (0.22° on a rotated pole grid) respectively. ~~As presented in Figure 1, the RCM 135 grid boxes are rotated by 0.22°.~~

7. "removing the RCM low precipitation" => any situations where you also create new wet days?
- (Reply) No, there are no cases where observations have more wet days than the RCM output at low precipitation. We have added this in the manuscript.

8. I suggest to delete the left panel in Figure 2.
- (Reply) Deleted as per the suggestion.

9. I suggest to move the model description, at least the detailed equation, to an Annex. Only describe those model components that are crucial for your bias correction argument.
- (Reply) Revised as per the suggestion.

10. This (Figure 6. Structure of the IHACRES model) can be retained in the main text
- (Reply) Revised as per the suggestion.

11. This (Table 1. Parameters in the IHACRES model) could go to the Annex.
- (Reply) Revised as per the suggestion.

12. "The observed precipitation and the two different bias corrected precipitation data from the conventional and proposed bias correction methods are randomly resampled to estimate the spread of the simulated flow ensembles." => This I don't understand. Why do you need a resampling to produce flow calculations?
- (Reply) To clarify, we have corrected the text as follows.

Second, the observed precipitation and the two different bias corrected precipitation time series from the conventional and proposed bias correction methods are constructed to compare the spread of simulated flow ensembles through the IHACRES model.

13. "Third, the optimised parameters and the ***precipitation time series*** are then used to simulate daily flow ensembles." => For me the ensemble that you produce consists of the 11 HadRCM simulations, performed with 2 different bias correction schemes. So I would mention "ensemble of precipitation forcings" here.

- (Reply) Corrected as per the suggestion.

14. Figure 7(b) => Delete 2nd word "parameter" from the legend.

- (Reply) Corrected as per the suggestion. Figure 9(b) and 10(b) are corrected as well.

15. Figure 10 => It is more clear when you plot panel (b) and (c) next to each other.

- (Reply) Corrected as per the suggestion (Figure 11 in the revised manuscript).

16. "Figure 11(b) shows the range of monthly spread" => It is unclear that this is actually the difference between the two lines you plot in panel (a).

- (Reply) To clarify we have revised the main text as follows (It is Figure 12 in the revised manuscript). Regarding the Figure caption, please refer to the reply to the 17th comment.
  (Before) Figure 12(b) shows the range of monthly spread.
  (After) Figure 12(b) shows the range of monthly spread i.e. the difference between the two lines in Panel (a).

17. "Figure 11. The spread of monthly mean flow for the period 1961-1990 derived from the precipitation ensembles." => Please describe panels (a), (b) and (c) in more detail.

- (Reply) We have revised the caption of Figure 12 (in the revised manuscript) as follows.

  (a) The range of monthly mean flow simulated from the precipitation ensembles for the period 1961-1990 (5-95 percentile spread); (b) The range of monthly spread in Panel (a); (c) Annual average value of the range in Panel (b).

18. Figure 11. Caption => You mean: range? Is it the absolute range, or is it a certain quantile?

- (Reply) Please refer to the reply to the 17th comment.

19. "Since different members are the outputs from different parameterisations, they would have different biases and be considered as ***independent (although not totally independent)*** from other ensembles."
    => This is a very vague formulation; please make clear what you mean by "independent".

- (Reply) We agree that 'independent' could confuse readers. To clarify this we have revised the text as follows.
- (Revised) Since different members are the outputs from different parameterisations, they would have different aspects of biases. Therefore, although the evidence is not conclusive, it is more reasonable to

consider each ensemble member on its own and conduct the bias correction separately for each member rather than pooling all of the members in the bias correction procedure.

20. Figure 13 => Personally I don't think this figure makes your paper more clear and attractive. I suggest to delete it, as it gives more noise and speculation.
- (Reply) We have deleted it.

21. "Figure 13 describes why the bias corrected members should originate from within the bounds of the natural variability of the observation." => You mean: the uncorrected RCM runs? Or do you mean: after bias correction they END UP (instead of ORIGINATE) within the boundaries?
- (Reply) We have deleted the figure.

22. "However, the proposed scheme is just one of the necessity conditions to assess the RCM ensembles and a comprehensive scheme including more conditions needs to be further developed."
=> What do you mean "more conditions"?
- (Reply) There could be more conditions in assessing the RCM ensembles which we do not know at this stage. To clarify this, we have revised the sentence as follows.
- (Before) a comprehensive scheme including more conditions needs to be further developed.
- (After) a comprehensive scheme including more conditions, if any, needs to be further developed.

23. Please acknowledge the reviewers.
- (Reply) We have acknowledged the reviewers as follows.
- Finally, we are grateful to Editor Bart van den Hurk, Reviewer Christiana S Photiadou and an anonymous reviewer for their valuable comments and suggestions on the manuscript.

24. The following corrections have been made as per the suggestions. The line numbers here are those in the annotated manuscript.
- L97: as an → to the
- L99: for the whole → to all
- L158: found in → matched with an equivalent cumulative probability in
- L212: are → is
- L294: delete "since the bias correction has been done on monthly basis" and "then" → deleted
- L294: under → for the two
- L308-310: delete → deleted
- L363: simulated precipitation → model simulations
- L370: realistic → variable
- L374: spread → range
- L376: delete "spread" → deleted
- L452: necessity → necessary