Responses to Reviewer #2

We are grateful to reviewer #2 for his/her constructive comments and suggestions, which are helpful to improve the quality of our manuscript. And we will make great efforts to address all the comments, with the details explained as follows.

General comment: The paper applies a Bayesian framework to assess the uncertainty in future projections of drought hazards in China. From reading the abstract and introduction, I am still unsure as too exactly what is the motivation of the study, and what the main hypothesis is. This needs to be made clearer to the reader. The main findings are that the frequency of drought event decreases, but the severity of the events increase. Further they claim that the method can detect and correct for uncertainties in the underlying RCM. I do find the results interesting, and the finding that the drought severity increases is important, but it needs to be further tested for significance and robustness. I also think that the paper reads more like a technical than a scientific paper, and this needs to be improved. I think that the paper needs a substantial revision before it can be published.

Response: To address the reviewer's comment, we will clarify the motivation and the main hypothesis in the Introduction section. To test the significance and robustness of the BMA technique, we will compare the AEM- and BMA-simulated hydroclimatic regimes and droughts through quantitative evaluation indices. In addition, we will add more explanations on the results and discussions on the underlying mechanism in the revised version of the manuscript, in attempt to provide new insights into the climate-induced drought risks.

Comment #1: The paper is very technical, and the main point is to apply the Bayesian framework on drought estimations on regional climate model output. I would argue that the uncertainties in both using regional climate modelling and the SPEI are highly uncertain in their underlying assumptions, so trying to correct any errors in the output is almost impossible. It is also a risk that sharpening the results leads to a wrong conclusion, since the methods applied might be completely off. A thorough testing of the methodology on for example reanalysis data would be one way of testing the robustness. The use of SPEI puzzles me, since I do not see how the values can be derived. They are not what is usually found in literature. Usually negative values denote drought conditions, and the values are within a few standard deviations, values outside the range +/-2 is usually considered very wet/dry. Please explain this more.

Response: To address the reviewer's comment, we will use a split-sample test to evaluate the robustness of the methodology in simulating hydroclimatic regimes and droughts. In addition, we will provide more details on the drought index (i.e., SPEI) in the revised version of the manuscript.

Comment #2: Drought indices do not necessarily indicate drought conditions. Each location has their own sensitivity, so the risk of drought should be considered in a study to say something about a severe event, this is captured in a comment below.

Response: We agree that drought indices do not necessarily indicate drought conditions. According to the reviewer's comment, we will change the term "drought risk" to "drought hazard".

Comment #3: Does the Bayesian framework bring any new light on this topic? What do the raw RCM results say in terms of increase severe drought indices? Are those results too muddled to draw any real conclusions from them? Is that because of a lack of precision in the modelling or because the uncertainties are so vast that it is really difficult to quantify these changes? Please provide these details to motivate why this methodology is necessary for this specific problem. The presentation can be improved. As a reader I am still confused as what the motivation really is behind the study. The numbers of figures are too many and not always relevant. I would like to see more skill assessments and less descriptive figures.

Response: According to the reviewer's comment, we will assess the hydroclimatic regimes and droughts generated from the Bayesian framework by comparing against the previous methods, such as the ensemble mean simulation. We will also discuss the deficiency of the previous methods used to simulate droughts.

Comment #4: L70 Calculating the drought return period does not necessary quantify the drought risk, rather the drought hazard. In order to get the risk you need to also take into consideration the impact.

Response: According to the reviewer's comment, we will change the term "drought risk" to "drought hazard".

Comment #5: Figure 1 contains a lot of information and is very difficult to decipher without first having read the paper. I know you are trying to show the work flow, but I would suggest to remake the figure to make it more schematic, with a few examples illustrating the stages.

Response: According to the reviewer's comment, Figure 1 will be revised for better presenting the workflow.

Comment #6: Figure 3. Please include the RCM names in the figure caption. Also, please use the same scale on the x-axis to make the interpretation easier.

Response: Figure 3 will be revised according to the reviewer's suggestion.

Comment #7: L245-262 and figure 6. In the comparison between AEM and BMA in terms of bias and correlation, the authors suggests that BMA outperforms AEM. I do not find that so evident. The correlation generally increases, but the standard deviation is generally worsened. This is noted by the authors, but I would suggest a more thorough analysis of this.

Response: To address the reviewer's comment, we will provide a through comparison on the performance of the AEM and BMA simulations, and then we will revise the irrelevant statements.

Comment #8: L 263-278, Figures 7-8. I would agree with eh authors that PET is generally improved with BMA, but I do not find the results for precipitation do so. For many regions BMA seems to smoothen the annual cycle of precipitation by increasing the winter precipitation and decreasing the summer precip. Also, the spread is reduced to a very thin band, which is not really what you want, This is not decreasing the uncertainty, this is being to over-confident with your

technique.

Response: We agree that the BMA technique does not lead to an all-round enhancement upon the AEM approach in simulating precipitation. Consequently, we will use a more appropriate BMA method, such as the copula-based BMA technique proposed by Madadgar and Moradkhani (2014), to improve the reliability of precipitation simulations.

Reference:

Madadgar, S., and Moradkhani, H.: Improved Bayesian multimodeling: Integration of copulas and Bayesian model averaging, Water Resour. Res., 50, 9586–9603, doi:10.1002/2014WR015965, 2014.

Comment #9: Figure 9, the bars under each section (a-f,g-l) only needs to be given once

Response: Figure 9 will be revised according to the reviewer's suggestion.

Comment #10: Section 3.2 L280-302. You show here the BMA-derived projections, but how do they differ from just using the ensemble mean of the RCM? Are the projected changes significantly different, and if so, why is that?

Response: To address the reviewer's comment, we will compare the drought projections generated from the BMA technique and the ensemble mean simulation, and then we will also analyze the underlying reason on the difference.

Comment #11: L330-339. I do not understand figure 13. I see the point that you want to compare the copula estimations and their uncertainty, but I do not understand how you can get the SPEI values and their estimated return period. An SPEI value of SPEI of 17.5 is impossible. The value of SPEI. or any standardised index, translates to the number of standard deviations away from the mean, where 1 is one standard deviation away from normal. Or am I misinterpreting the figure? If so, please help me interpret the figure. Same goes for figure 14

Response: We regret for the unclear statement on the calculation of drought characteristics. For better clarification, we will provide detailed statement on the calculation of drought characteristics and the return period. We will also provide more detailed descriptions on Figures 13 and 14.

Comment #12: L389. The authors claims that the reliability of the RCM output is improved, but I see no sign of that ever tested? I would suggest to add reliability diagrams to test this hypothesis properly.

Response: To address the reviewer's comment, we will use reliability diagrams and quantitative indices to evaluate the reliability of the BMA-based climate simulations.

Comment #13: Conclusions. The authors claim that BMA can "successfully correct" the errors. I disagree with this categorical statement. Some aspects are improved, but others are worsened. This should be clearer.

Response: According to the reviewer's suggestion, the section of Conclusions will be revised for better clarification.