Responses to the comments from Anonymous Referee #1

We are very grateful to the reviewer for the positive and careful review. The thoughtful comments have helped improve the manuscript. The reviewer's comments are italicized and marked in blue, and our responses immediately follow.

The manuscript by Jiao and Yuan assessed the possible changes of drought characteristics (frequency, duration and severity) under future climate at Wudinghe watershed in the semiarid region of China, which is one of the largest sub-basins of the Yellow River basin. The content generally falls into the interests of HESS and its broad audience. Overall, the technical framework is well designed and the manuscript is in good shape for publication. I suggest a minor revision for the authors to address my following concerns.

Response: We would like to thank the positive comments from the reviewer. Please see our responses below.

First, I found some critical details are missing in section 2 and 3 of this manuscript. Most importantly, there is no details on temporal disaggregation of the GCM-based Ta and Prec. Also, there is no information about other input variables for the CLM-GBHM model. Moreover, the performance of CLM-GBHM model in reproducing the historical streamflow is largely unknown, though there is some validation work in previous works (Jiao et al., 2017; Sheng et al., 2018). To make the future projection more convincing, the authors should first demonstrate the model performance in the whole baseline period (1986-2005) considering that Jiao et al. (2017) only showed the model validation results during 1964 to 1969, which is out of the baseline period here.

Response to R1C1: Thanks for the comments and advices. The first comment on "temporal disaggregation" is further explained in **Response to R1C5**, and response to second comment about "other input variables" for the CLM-GBHM could be found in **Response to R1C6.** As for model performance, we have now compared the simulated and observed standardized streamflow index (SSI) during the baseline period (1986-2005) as follows:

"Therefore, we took 1986-2005 as the baseline period. Monthly standardized streamflow index (SSI) simulations from CLM-GBHM were compared to the observed records during the baseline period, and the model performed well with a correlative coefficient of 0.53 ($p<0.01$)."

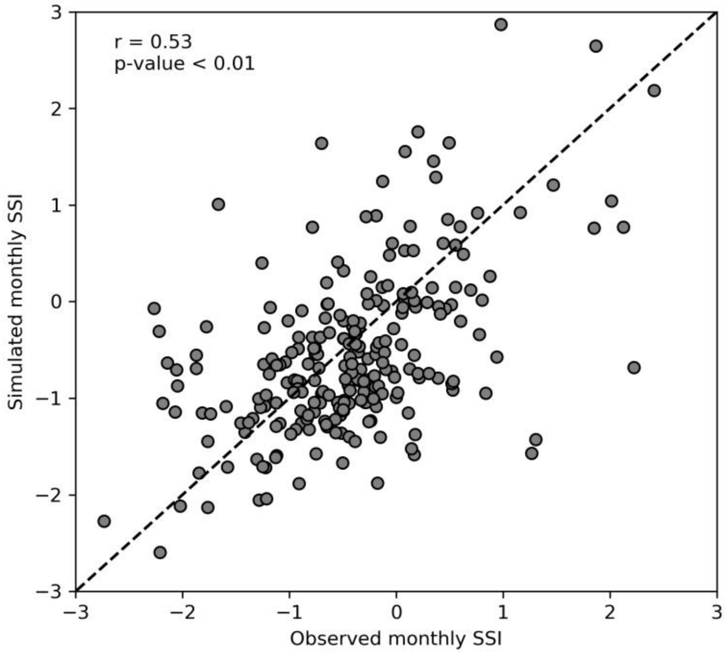

Figure R1: Model verification for monthly standard streamflow indices during baseline period (1986-2005)

Second, the uncertainty separation framework is valid for GCM outputs. However, for streamflow and drought frequency, the model should be "GCM+CLM-GBHM". If the error propagation in the CLM-GBHM is totally linear (which is the assumptions of the current manuscript), then the uncertainty contribution ratios for "GCM+CLM-GBHM" should be the same with those for the "GCM". Otherwise, they may be different.

Response to R1C2: We agree with the reviewer. Because of the complex interaction between biosphere and hydrosphere, the land surface model CLM-GBHM has a nonlinear error propagation. We have revised the legend in Figures 8c and 8d from "Climate models" to "Climate and land surface models".

Other minor comments:

P6L91: please specify the time range for the "long-term annual mean…"

Response to R1C3: Thanks for the comments. We have specified the time range and revised the manuscript as follows:

"It has a semiarid climate with long-term (1956-2010) annual mean precipitation of 356 mm and runoff of 39 mm, resulting in a runoff coefficient of 0.11 (Jiao et al., 2017)"

P7L104: please justify the choice of "eight" GCMs. Do you have any criteria for this *selection? Will the selection affect the later analysis?*

Response to R1C4: Thanks for the advices. We chose those CMIP5 GCMs with publicly accessible daily precipitation and air temperature simulations under all four RCP scenarios, and finally got eight GCMs in this study. We have clarified as follows:

"In this study, we chose eight CMIP5 GCMs for historical (1961-2005) and future (2006-2099) drought changing analysis, as they provided daily simulations under all four RCP scenarios (i.e., RCP2.6/4.5/6.0/8.5)."

P7L117-119: this temporal downscaling should be elaborated in more details.

Response to R1C5: Thanks for your comments. In this study, all CMIP5 simulation data were collected at daily scale, but the bias correction was performed at monthly scale. After that, new daily precipitation series were generated based on the ratio of new and old monthly mean results, and daily temperature data were based on the difference between new and old monthly means:

$$P_{d,no-bias} = \left(\frac{P_{m,no-bias}}{P_{m,with-bias}}\right) P_{d,with-bias}$$

$$T_{d,no-bias} = \left(T_{m,no-bias} - T_{m,with-bias}\right) + T_{d,with-bias}$$

where *P* and *T* represent precipitation and temperature, subscripts *m*, *d*, *with-bias*, *no-bias* represent monthly mean value, daily value, value with bias and value after bias correction, respectively. After that, CRUNCEP 6-hourly climate dataset (https://svn-ccsm-inputdata.cgd.ucar.edu/trunk/inputdata/atm/datm7/) during 1959-2005 were collected for temporal downscaling from daily to 6-hourly scales by using similar method:

$$P_{6h,no-bias} = \left(\frac{P_{d,no-bias}}{P_{d,CRUNCEP}}\right) P_{6h,CRUNCEP}$$

$$T_{6h,no-bias} = \left(T_{d,no-bias} - T_{d,CRUNCEP}\right) + T_{6h,CRUNCEP}$$

where subscripts *6h* and *CRUNCEP* represent 6-hourly value and value from CRUNCEP data. We have modified the manuscript as follows:

"…to fit their cumulative density functions to observed ones based on monthly mean value. Bias-corrected daily precipitation and temperature were then further temporally disaggregated to a 6-hours interval based on their intraday distribution from CRUNCEP (https://svn-ccsm-inputdata.cgd.ucar.edu/trunk/inputdata/atm/datm7/)

6-hourly dataset for driving land surface hydrological model.”

Section 3.1: what's the input variables needed for CLM-GBHM model? Besides Ta and Prec, there should be some other variables. How would you deal with those other variables and what's the data sources?

Response to R1C6: Thanks for the comments. The input climate forcing variables used by CLM-GBHM include precipitation, air temperature, incident solar radiation, air pressure, specific humidity and wind speed. We took CRUNCEP data during 1959-2005 (47 years) to get these variables needed for simulation. Historical (1961-2005, 45 years) variables were directly taken from CRUNCEP data; future (2006-2099, 94 years) variables were generated by looping the CRUNCEP data twice. We have specified this by adding the follows to the end of Section 2:

“Other 6-hourly climate forcings, i.e. incident solar radiation, air pressure, specific humidity and wind speed, were directly taken from CRUNCEP data.”

P8L132: Why did you choose to use the monthly LAI of 1982 for all the experiments? Please justify this. Would use the historical climatology of LAI (say from 1986 to 2005) be more reasonable here?

Response to R1C7: Thanks for the comments. Our previous work (Yuan et al., 2018, WRR) considered the vegetation dynamics in this area, and showed that vegetation variation contributed only a small proportion to historical changes in streamflow and extremes. As vegetation dynamics is not the main concern of our paper and future vegetation variation is unknown, there would be further work on this topic, while here we simply fixed LAI to the value in 1982.