Replies to Referee #1

Transferability of climate simulation uncertainty to hydrological climate change impacts

Hui-Min Wang, Jie Chen, Alex J. Cannon, Chong-Yu Xu, Hua Chen

We sincerely appreciate the referee's comments and suggestions on the manuscript. All suggestions are helpful to improve this manuscript. We have carefully studied, considered and responded to all comments point-by-point as follows. For clarity, all comments are given in italics and responses are given in plain text. The manuscript will be modified accordingly.

The manuscript is about the transferability of the climate model uncertainties, introduced by the selection of climate models, to hydrological impacts. To this end, two envelope-based selection methods, K-means clustering and the Katsavounidis-Kuo-Zhang (KKZ) method, are used to select subsets from an ensemble of 50 climate models over two watersheds with different climate characteristics. The transferability of the climate model uncertainties is evaluated by comparing uncertainty coverage between 31 climate variables and 16 hydrological variables that are simulated by the hydrological model GR4J. In addition, also the importance of choosing climate variables properly while selecting subsets is investigated by in- and excluding temperature variables. The manuscript is well structured and written. The manuscript covers a topic that is original and novel, and might interest a large amount of readers, including climate scientists and hydrologists. To my opinion, this manuscript needs some minor revisions. I have added a few comments/suggestions that need to be addressed before acceptance.

We appreciate that the reviewer is in favor of the content of this research. All the comments and suggestions have been replied to below and will be addressed in the revision.

General Comments

I would rephrase the title a bit. In the title, the authors are referring to the transferability of climate simulation (model?) uncertainty to "hydrological climate change impacts", whereas in the Abstract and other parts of the manuscript the authors write about the transferability to "hydrological impacts". I would change the "hydrological climate change impact" into "hydrological impacts". In this way, the authors can put more emphasis on the transferability of uncertainties to "hydrological impacts" specifically, and the title has a better connection with the Abstract and manuscript or vice versa.

Thank you for the suggestion on the title. The "hydrological climate change impacts" will be changed to "hydrological impacts" in the revised manuscript. With regard to the use of "climate simulation" instead of "climate model", although global climate models (GCMs) are the main uncertainty source considered, this research also includes uncertainty related to future emission scenarios (i.e. RCP4.5 and RCP8.5). Thus, we think that "climate simulation" is more appropriate than "climate model" in the title.

Following reviewer's suggestion, the title will be changed to "Transferability of climate simulation uncertainty to hydrological impacts" in the revised manuscript.

Two watersheds with different climate characteristics are selected for this study. It would be good to spend some text in the Discussion explaining what the potential effects of transferability are in other climatic regions, such as high-mountain regions.

Thanks for the comment. We agree with the reviewer that it is necessary to further discuss the transferability in other climatic regions.

For example, for high-mountain regions, precipitation may be influenced by complex topography and snowmelt often has the greatest contribution to runoff. Therefore, it is recommended that climate variables related to orographic precipitation and the evolution of snowpack be included in the selection process. Additionally, for arid regions, Hortonian overland flow may be the predominant runoff mechanism, and thus precipitation variables that are capable of describing the intensity of precipitation events may need to be stressed in the selection of subsets.

These points will be added to the Discussion section of the revised manuscript.

Specific Comments

1. Introduction; L4-9: the authors indicate that the selection methods inherit the potential flaws of the past-performance approach, when the emphasis is on model performance. What are the potential flaws of the past-performance approach? Combining envelope coverage criteria and past-performance would, to my opinion, be better since not only the models are selected to represent a full range of climate conditions, but also are tested in their performance to simulate regional (historical) climate characteristics, especially in those regions where, for instance, monsoon systems prevail.

Thank you for the comment. In this sentence, the potential flaws are meant as "In general, the assumption that models with good performance over the near-past provide more realistic climate change signals is questionable (Knutti et al., 2010; Reifen and Toumi, 2009), and the metrics commonly used to evaluate model performance are often manually defined based on the fields of interest, which leads to substantial subjectivity within the weighting process (Mendlik and Gobiet, 2016)". Specifically, the best performing models may not produce the most realistic climate change signal in the future. In addition, the ranking or weighting GCMs is highly dependent on the definition of metrics. Thus, the past-performance approach may lead to subjectivity in the selection of climate model simulations.

With regards to combining envelope coverage criteria and past-performance criteria, on one hand, we actually wanted to say (but failed to say it clear enough) that many approaches combining both criteria put more emphasis on the past-performance criteria. These methods may inherit the potential flaws of the past-performance approach. On the other hand, we agree with the reviewer that combining envelope coverage criteria and past-performance criteria may be better at selecting climate simulations for impact studies. In other words, it may be more reasonable to remove unrealistic models rather than keep models with “best performance”. This point has been stated in the 3rd paragraph of the Introduction [Page 2, Lines 26-27] and in the last paragraph of the Discussion section [Pages 15, Lines 9-11]. This question deserves further investigation.

All related information will be clarified in the revised manuscript.

> *1. Introduction; L20-25: To my opinion, the number of variables that is chosen for a selection approach depend on what the scope is of the study. If a study has only a focus on projecting changes in water availability or changes in the water balance it would, to my opinion, not be necessary to take indices into account that represent climatic extremes, whereas a study with a focus on hydrological extremes needs to include these indices in the selection approach. Therefore, the authors need to elaborate more on why a certain number of variables should be selected or not.*

Thanks for this comment. We agree with the reviewer that there may be not necessary to take so many climate extreme indices into consideration when a study does not focus on hydrological extremes. However, it is difficult to determine a one-to-one linkage between climate and hydrological variables due to the non-linearity of hydrological responses. Only a few climate variables may be not enough to describe climate conditions which have great influence on hydrological processes. The main objective of this study is to investigate the transferability of climate simulation uncertainty to hydrological impacts. The results show that including climate extreme indices improves the transferability of climate simulation uncertainty, compared with the results of Chen et al. (2016). Therefore, climate conditions described by extreme indices are found to be important for the transferability of climate uncertainty.

These issues will be discussed in the revised manuscript.

> *2.2.1 Climate Simulations; L9: Why did the authors select 50 models? The authors mentioned before that the CMIP5 archive includes 61 models. Is there a reason why the other 11 models are excluded from the selection approaches? In addition, each climate model has one or more ensemble member. Did the authors select the first ensemble member or did they select random ensemble members? The authors need to include this information in the method description, for instance by adding extra information to Table 1.*

Thanks for your comments. We actually employed 26 GCMs from the CMIP5 ensemble (Table 1) with simulations based on two Representative Concentration Pathways (RCP4.5 and RCP 8.5), with the exception of CMCC-CESM, which only used RCP8.5, and MRI-ESM1, which only used RCP4.5. On the whole, 50 climate simulations were used. Some GCMs in the CMIP5 ensemble were excluded due to lack of relevant variables (e.g., daily outputs that are necessary to drive the hydrological model) or lack of temporal coverage (e.g., the reference or future periods used in this study). In addition, GCMs employed in this study are consistent with Chen et al. (2016) to make the two studies more comparable.

In addition, only the first ensemble member of each GCM was used; this information will be clarified in the Data section.

2.2.2 Observations; L17: Where are the data from the 100 rain gauges, 8 temperature gauges, and 1 streamflow gauge obtained? The authors have to include some references to the sources where they obtained the meteorological and discharge data.

Thanks for the suggestion to include references to the sources of observational data. Observations in the Xiangjiang watershed are the same as those used by Zeng et al. (2016) and Xu et al. (2013) and were provided by the Changjiang Water Resources Commission. This information will be clarified in the Data section.

3.2 Generation of Climate Scenarios: Why did the authors use the DS method to downscale GCMs and not a method such as the Advanced Delta Change approach or the Quantile Mapping approach that also take changes in extremes into consideration? It might be interesting to discuss potential uncertainties that are introduced by the downscaling approaches in the Discussion.

Thanks for the comment. We failed to describe clear enough that Daily Scaling (DS) method is an advanced delta change approach combining delta change and quantile mapping methods. The rationale behind using a change factor method instead of a bias correction method is that climate change signals may be modified by some forms of bias correction (e.g., see Cannon et al., 2015). Still, in order to investigate the influence of alternative bias correction/downscaling approaches, we have also used the Quantile Mapping (QM) approach to post-process the CMIP5 GCMs. Figure R1 shows the results of average PSC in this case. For the Xiangjiang watershed, temperature variables were excluded in the process of selection, while they were included for the Manicouagan 5 watershed. Compared with the results calculated by the DS method, the overall characters of the results are mostly the same. The choice of the downscaling method has little influence on the conclusions of this study. However, to highlight the potential sensitivity of results to different downscaling methods, this information will be added to the Discussion section of the revised manuscript.

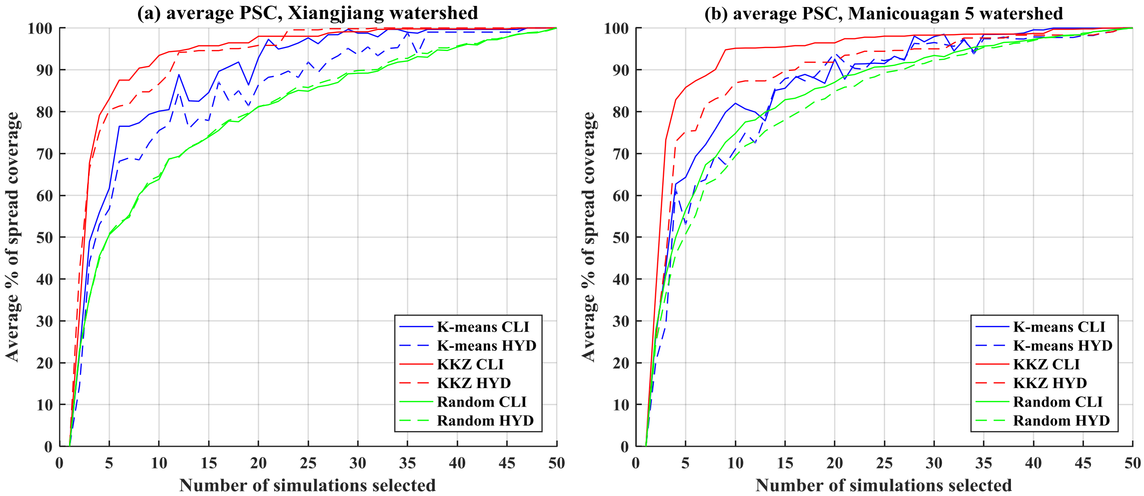

Figure R1: The average PSC for three different selection methods over two watersheds when using the QM approach (CLI = climate variables and HYD = hydrological variables).

> *3.3.1 Hydrological Modelling; L15-21: I would recommend replacing and to discuss this part in more detail in the Results Section, for instance under a separate subsection "Calibration and Validation"*

We agree with the reviewer. The calibration and validation of the hydrological model will be presented in the Results section as follows.

The basin-averaged daily minimum and maximum temperature and precipitation in calibration and validation periods, as shown in Table R1, were used to calibrate and validate the GR4J-6 model over the two watersheds. Model parameters were obtained by the shuffled complex evolution optimization (Duan et al., 1992) based on the objective to maximize Nash-Sutcliffe Efficiency (NSE) (Nash and Sutcliffe, 1970). The optimally chosen sets of parameters yield a NSE between 0.87 and 0.93 over both watersheds. The calibrated GR4J-6 also yields small relative errors of water balance with values between -0.3% and 5.4% for calibration and validation, which demonstrate the applicability of the calibrated model over both watersheds (Table R1). Figure R2 presents the mean daily hydrographs (average discharge of each calendar day across the years) for the calibration and validation periods over two watersheds. The calibrated GR4J-6 has good performance in most of the year, with the exception that discharge is overestimated in the winter for the Manicouagan 5 watershed. In addition, since snow accumulation and snowmelt processes are not important in the Xiangjiang watershed, the GR4J model (excluding snow module) was also calibrated in this watershed. Results showed that there was little difference between the calibrated GR4J and GR4J-6 models, and thus the presence of the CemaNeige snow module would not influence the performance of GR4J-6 in the rainfall-characterized Xiangjiang watershed (Table R1).

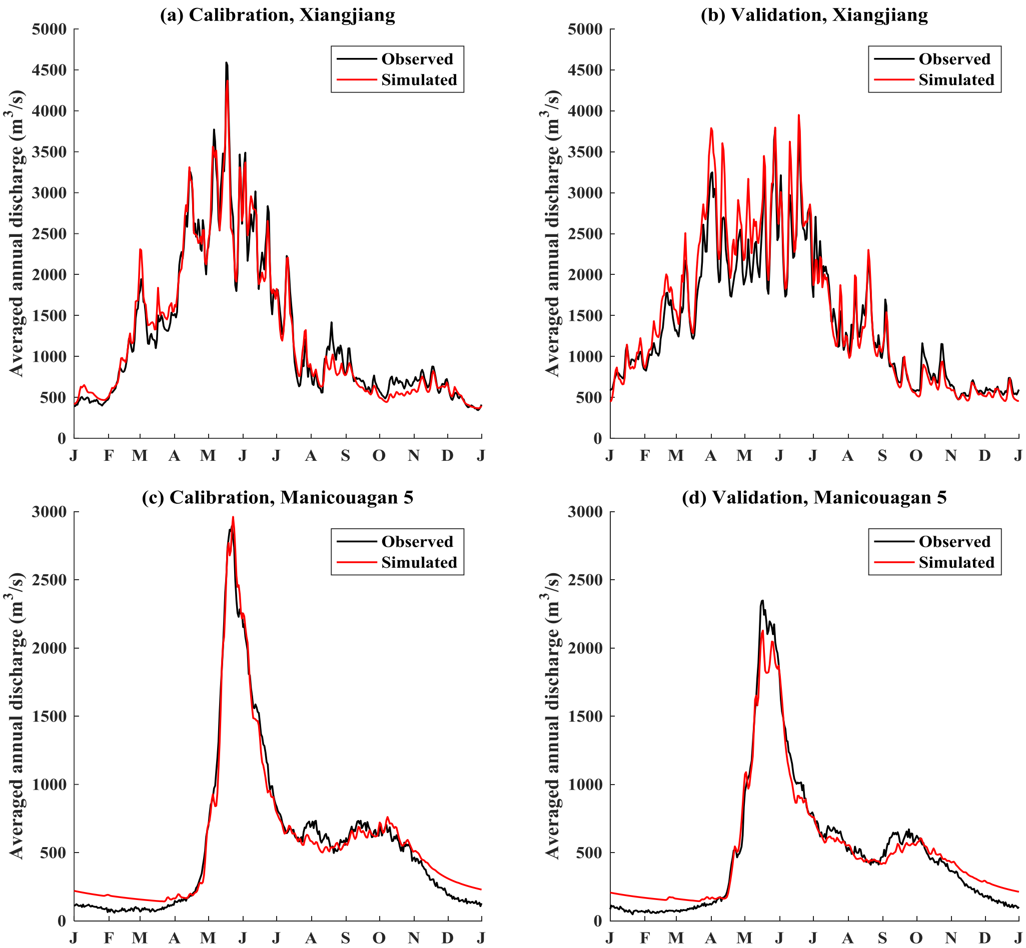

Figure R2: Observed and simulated mean hydrographs for (a, c) calibration and (b, d) validation periods over the (a, b) Xiangjiang and (c, d) Manicouagan 5 watersheds.

Table R1: Nash-Sutcliffe Efficiencies (NSE) and relative errors of water balance (RE) of hydrological models in the calibration and validation over two watersheds

Country	Watershed name	Area (km^2)	Hydrological Model	Calibration period	NSE calibration	RE calibration	Validation period	NSE validation	RE validation
China	Xiangjiang	52150	GR4J-6	1975-1987	0.912	-0.3%	1988-2000	0.871	5.4%
			GR4J	1975-1987	0.912	-0.2%	1988-2000	0.872	5.5%
Canada	Manicouagan 5	24610	GR4J-6	1970-1979	0.926	3.8%	1980-1989	0.881	2.7%

3.4 Data Analysis: It would be good to an additional sentence on what it means when a PSC is high or low. For the reader, it might be more difficult to image the meaning of a high or low PSC.

Thanks for your suggestion. A higher PSC means that the selected subset covers a larger uncertainty range. This will be clarified in the revised manuscript.

Figure 1: The longitude axes are given, but the latitude axes are missing. Further I think the figure does not contain a lot of information. I would add a digital elevation model or another topographic/geographic info to give the reader more valuable information on the characteristics of the catchments. In addition, I would add the positions of the discharge gauging stations used for the calibration/validation. Finally, I recommend making inlets,

including the catchment maps, larger.

We agree with the reviewer. Figure R3 will be updated in the revised manuscript.

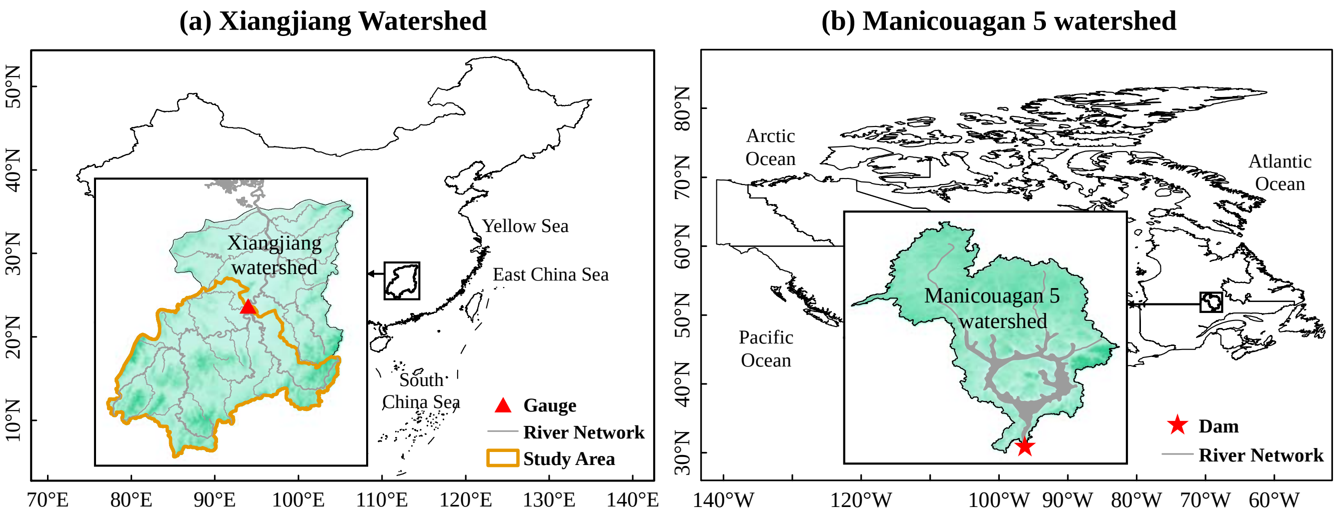

Figure R3: Location maps of the (a) Xiangjiang and (b) Manicouagan 5 watersheds (The study area in the Xiangjiang watershed is one of its sub-basins as the orange boundary shows).

Technical Comments

Abstract; L16: ". . .the importance of choosing climate variables properly while selecting subsets. . ." instead of ". . .the importance of properly choosing climate variables in selecting subsets. . ."

This sentence will be revised as suggested.

References

Cannon, A. J., Sobie, S. R., and Murdock, T. Q.: Bias Correction of GCM Precipitation by Quantile Mapping: How Well Do Methods Preserve Changes in Quantiles and Extremes?, Journal of Climate, 28, 6938-6959, https://doi.org/10.1175/JCLI-D-14-00754.1, 2015.

Chen, J., Brissette, F. P., and Lucas-Picher, P.: Transferability of optimally-selected climate models in the quantification of climate change impacts on hydrology, Climate Dynamics, 47, 3359-3372, https://doi.org/10.1007/s00382-016-3030-x, 2016.

Duan, Q., Sorooshian, S., and Gupta, V.: Effective and efficient global optimization for conceptual rainfall-runoff models, Water Resources Research, 28, 1015-1031, https://doi.org/10.1029/91WR02985, 1992.

Knutti, R., Furrer, R., Tebaldi, C., Cermak, J., and Meehl, G. A.: Challenges in Combining Projections from Multiple Climate Models, Journal of Climate, 23, 2739-2758, https://doi.org/10.1175/2009jcli3361.1, 2010.

Mendlik, T., and Gobiet, A.: Selecting climate simulations for impact studies based on multivariate patterns of climate change, Climatic Change, 135, 381-393, https://doi.org/10.1007/s10584-015-1582-0, 2016.

Nash, J. E., and Sutcliffe, J. V.: River flow forecasting through conceptual models part I — A discussion of principles, Journal of Hydrology, 10, 282-290, https://doi.org/10.1016/0022-1694(70)90255-6, 1970.

Reifen, C., and Toumi, R.: Climate projections: Past performance no guarantee of future skill?, Geophysical Research Letters, 36, https://doi.org/10.1029/2009gl038082, 2009.

Xu, H., Xu, C.-Y., Chen, H., Zhang, Z., and Li, L.: Assessing the influence of rain gauge density and distribution on hydrological model performance in a humid region of China, Journal of Hydrology, 505, 1-12, https://doi.org/10.1016/j.jhydrol.2013.09.004, 2013.

Zeng, Q., Chen, H., Xu, C.-Y., Jie, M.-X., and Hou, Y.-K.: Feasibility and uncertainty of using conceptual rainfall-runoff models in design flood estimation, Hydrology Research, 47, 701-717, https://doi.org/10.2166/nh.2015.069, 2016.