Thank you for the opportunity to review this manuscript. This study deals with an important issue on uncertainty in hydrological models. The authors calibrated the VIC model using a multiobjective approach based on historical data. Sixteen sets of model parameters were randomly selected. Then the model with these 16 sets of parameters were applied to two future phases up to year 2100 with 18 different scenarios. The changes in system states were represented using ensemble means. It is an interesting study. However, it is not very clear what is the major contribution of this study considering the large number of studies on uncertainty in the literature. In addition, some major modeling steps/decisions are not described clearly or justified. Therefore, this reviewer recommends major revision.

Reply: We thank Reviewer 1 for the constructive comments and suggestions that will improve the quality of this work. We have carefully considered all suggestions and outlined a set of proposed revisions in the following response.

Detailed comments: 1. Uncertainty is important as pointed out by the authors. However, what is the major contribution of this paper considering there are a large number of studies dealing with uncertainty in hydrological models? One of the objectives of the study is to "quantify the uncertainty resultant from model parameters to projections of hydrologic flux and state variables". Has this never been done before? If so, please support with evidence in the introduction section.

Reply: We agree that there are a large number of studies dealing with uncertainty in hydrologic models. That said, there are considerably fewer studies that focus specifically on the role of parametric uncertainty in the context of future climate change projections, which is the focus of this paper. We make sure to more clearly articulate this distinction in the revised manuscript. In this context, Dobler et al. (2012) quantified uncertainties resulting from global or regional climate models, bias-correction method, and hydrological model parameterizations in an Alpine watershed of Austria. They concluded that hydrological model parameterization was the least important source of uncertainty. Other studies that have compared multiple uncertainty sources (Addor et al. 2014; Yuan et al. 2017; Joseph et al. 2018) have largely downplayed the importance of model parameterizations. Another study in India went as far as concluding that hydrologic model parametric uncertainty is negligible relative to meteorological uncertainty, (Joseph et al. 2018). If that was indeed true, then it would be justifiable to use any uncalibrated hydrological model to project future hydrological fluxes and state variables. However, a recent study demonstrated that while uncertainty of GCM projections was the dominant source for faster components of hydrologic response like surface runoff, the uncertainty of hydrological model parameterization was found to be a significant source of uncertainty, particularly for slow response components (Her et al., 2019). This finding indicates that calibration of hydrological models is still important for projecting some important hydrological variables. Another study also demonstrated that model parameters could be of major importance for projecting changes in water-quality (Steffens et al., 2014).

Most relevant to this manuscript, a few studies have investigated how hydrologic modeling decisions can affect future hydrological projections (e.g. Mendoza et al., 2016; Seiller et al., 2017). These studies considered the impacts of the selection of hydrological model, model structure, and parameter sets over a number of river basins. However, what remains unclear, and the focus of this manuscript, is whether the impact of parameter set selection alone is large enough to impact the direction and magnitude of projected changes, and if so, how the magnitude of this impact would translate across different temporal scales. For example, we evaluate whether two model parameter sets that have both been calibrated can produce a

different sign in climate sensitivity, e.g. increasing- versus decreasing projected future streamflow. As a result, we focused this work on the contribution of model parameter uncertainty to future hydrological projections, including the contribution of parametric uncertainty at different time scales (annual, monthly, and daily) and for different hydrological variables. Relative to previous studies, we focus here on only a single river basin, but investigate more thoroughly the details of model parameter contributed uncertainty at fine time scales (i.e. daily) and the resultant affects on hydroclimatic extremes.

2. The context in which parameter uncertainty is assessed. Parameter uncertainty is only one contributor of uncertainties, among model uncertainty, input uncertainty, climate uncertainty, etc. Where does parameter uncertainty sit among all uncertainties?
Reply: We agree that parameter uncertainty is only one contributor of uncertainties (see references to Mendoza et al., 2016; Seiller et al., 2017, among others) and that its magnitude relative to other uncertainties can vary. In larger (global) domains, parameter uncertainty might have a smaller impact on future projections due to the potential for compensating errors to partially offset at large spatial scales (Elsner et al., 2014). However, at smaller scales with less potential for errors to cancel out, like the one presented in this manuscript, the impact of parameter uncertainty has the potential to be large and potentially dominant. In order to evaluate where parameter uncertainty sits among other uncertainties, we compare it with the uncertainty from future climate change scenarios and emphasize this distinction in the revision.

3. The authors emphasized uncertainty related to Climate Change in introduction. But ensemble mean across different future scenarios are used in the study, where uncertainty represented by the different future scenarios is lost.
Reply: We clarify this issue in the revision. Namely that we consider both the ensemble mean—which is widely used for future projections—as well as the ensemble spread as a way to quantify the uncertainty across future climate scenarios. In this way, we aim to preserve the uncertainty represented by different future scenarios. We consider 18 future scenarios and 16 parameter sets to make a total ensemble of 18*16=288 members. The spread of the 288 members is used as a measure of the total uncertainty. For each parameter set, there are 18 future scenarios and the spread of these 18 future scenarios is used to quantify the uncertainty of future scenarios associated with the parameter set. We will therefore obtain the median of the 18 future scenarios for each parameter set and finally get the range from the 16 median values (as for the 16 parameter sets). Following the same approach, we can quantify the uncertainty associated with the different parameter sets. We have clarified the calculation in the revision.

4. The authors selected a large number of performance measures. Why were they selected? If they are randomly selected with no justification, the authors are increasing calibration effort without additional benefit.
Reply: We selected the parameter sets with similar performance but located in different regions of the parameter space following the approaches used by Moriasi et al. (2007) and Demaria et al. (2007). Performance measures were selected for the following specific reasons: NSE is the most commonly used metric in hydrologic modeling and quantifies the overall performance, putting an emphasis on the seasonal cycle. In contrast, RMSE measures the error in the squared units of the simulated versus observed datasets and emphasizes high flows and outliers. PBIAS measures the average of the simulations compared to the observed datasets, evaluating the quality of the overall water balance. R^2 is the coefficient of determination between observed and simulated datasets. The ratio of standard deviations exclusively evaluates the variability of the model

relative to the observation and thus represents a valuable quality control. We now clarify this rationale further in the revision.

5. The authors stated that "The 16 best performing parameter sets were chosen randomly using the Borg MOEA framework." This statement is confusing. If the "best" parameter sets are selected, they much be selected based on some criteria rather than randomly. Did the authors mean they are selected randomly from the Pareto-optimal front obtained from multiobjective optimization using Borg? Then why are they selected randomly? Will another different set of parameters selected randomly lead to different results and conclusions?
Reply: We thank the reviewer for making this point. Yes, we chose the parameters from the Pareto optimal front, so as to represent non-dominated solutions. That is, parameter sets with similar performance but located in different regions of the parameter space were chosen. The Borg MOEA uses the epsilon non-dominance operator, which has the advantages of convergence and diversity with respect to approximating the true Pareto-optimal front over other MOEA. The epsilons represent the resolutions of the objective functions. Specifically, the epsilon-box dominance archive divides the objective space into hyper-boxes with side-length epsilon, so called epsilon-boxes (Hadka and Reed, 2013). The 16 parameter sets represent epsilon-box non-dominated solutions that were sampled from the full Pareto-optimal front such that the conclusions here are expected to be robust. The selecting principle and method have been clarified in the revision.

6. One of the conclusions in the manuscript is that future changes in system state are similar to those in the past (Section 5.3). However, the model was calibrated using historical data then applied to future scenarios. What is the implication of this approach to the conclusion?
Reply: I think you refer to L354-357. The implicit assumption here is that parameter sets calibrated during historical periods can be applied to future simulations. The chosen parameter sets give satisfactory results during validation period historically. The focus here is how uncertainties across 'calibrated' parameter sets will change between past conditions relative to future climate. We make this assumption clearer in the revised manuscript.

7. The authors claimed that "variability due to parameter uncertainty was up to 10 % annually and 26 % monthly under future climate change scenarios". Without understanding the whole picture of uncertainty, it will be difficult to reach this conclusion.
Reply: Please see our response to #3 above and associated clarifications in the manuscript. These numbers represent the fraction (or percentage) of the uncertainty associated with parametric spread relative to the total (parametric + climate scenario) uncertainty.

Minor comments: 1. Description on the Borg MOEA between lines 148 and 152 is inaccurate.
Reply: We have updated the description of the Borg MOEA for accuracy.

2. Many terms are used in an ad hoc manner. For example, Borg is an optimization algorithm not a framework. Line 224: "Borg MOEA framework".
Reply: We correct and streamline the terminology and usage in the updated version.

3. Line 351 and Line 353: Remove "the" before "summer".
Reply: Will be corrected.

References

Addor, N., O. Rössler, N. Köplin, M. Huss, R. Weingartner, and J. Seibert (2014), Robust changes and sources of uncertainty in the projected hydrological regimes of Swiss catchments, Water Resour. Res., 50, doi:10.1002/2014WR015549.

Demaria, E. M., Nijssen, B., and Wagener, T. (2007) Monte Carlo sensitivity analysis of land surface parameters using the Variable Infiltration Capacity model. Journal of Geophysical Research, 112, D11113, doi:10.1029/2006JD007534.

Dobler, C., Hagemann, S., Wilby, R. L., and Stötter, J.: Quantifying different sources of uncertainty in hydrological projections in an Alpine watershed, Hydrol. Earth Syst. Sci., 16, 4343-4360, https://doi.org/10.5194/hess-16-4343-2012, 2012.

Elsner, M. M., S. Gangopadhyay, T. Pruitt, L. D. Brekke, N. Mizukami, and M. P. Clark (2014) How Does the Choice of Distributed Meteorological Data Affect Hydrologic Model Calibration and Streamflow Simulations? Journal of Hydrometeorology, 15(4):1384-1403, doi:10.1175/JHM-D-13-083.1.

Hadka, D., Reed, P. (2013) Borg: An Auto-Adaptive Many-Objective Evolutionary Computing Framework. Evolutionary Computation, 21, 231-259, doi: 10.1162/evco_a_00075.

Her, Y., Yoo, S. H., Cho, J., Hwang, S., Jeong, J., & Seong, C. (2019). Uncertainty in hydrological analysis of climate change: multi-parameter vs. multi-GCM ensemble predictions. *Scientific reports*, *9*(1), 4974.

Joseph, J., Ghosh, S., Pathak, A., & Sahai, A. K. (2018) Hydrologic impacts of climate change: Comparisons between hydrological parameter uncertainty and climate model uncertainty. Journal of Hydrology, 566, 1-22.

Mendoza, P. A., M. P. Clark, N. Mizukami, A. J. Newman, M. Barlage, E. D. Gutmann, and R. M. Rasmussen (2015) Effects of Hydrologic Model Choice and Calibration on the Portrayal of Climate Change Impacts. Journal of Hydrometeorology, 16:762-780, doi: 10.1175/JHM-D-14-0104.1.

Moriasi, D. N., Arnold, J. G., Van Liew, M. W., Bingner, R. L., Harmel, R. D., and Veith, T. L. (2007) Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. American Society of Agricultural and Biological Engineers, 50(3):885-900.

Steffens, K., Larsbo, M., Moeys, J., Kjellström, E., Jarvis, N., and Lewan, E.: Modelling pesticide leaching under climate change: parameter vs. climate input uncertainty, Hydrol. Earth Syst. Sci., 18, 479-491, https://doi.org/10.5194/hess-18-479-2014, 2014.

Younggu Her, Seung-Hwan Yoo, Jaepil Cho, Syewoon Hwang, Jaehak Jeong & Chounghyun Seong, Uncertainty in hydrological analysis of climate change: multi-parameter vs. multi-GCM ensemble predictions, Scientific Reports, 9, 4974 (2019).

Yuan, F., C. Zhao, Y. Jiang, L. Ren, H. Shan, L. Zhang, Y. Zhu, T. Chen, S. Jiang, X. Yang, H. Shen: Evaluation on uncertainty sources in projecting hydrological changes over the Xijiang River basin in South China, J. Hydrol., 554 (2017), pp. 434-450.