Interactive comment on "*Assessing Climate Change-Induced Flood Risk in the Conasauga River Watershed: An Application of Ensemble Hydrodynamic Inundation Modeling*" by Tigstu T. Dullo et al.

The authors would like to thank the reviewer for the insightful and constructive comments. We have reviewed the comments and provided our responses herein. The reviewer's comments are presented in first followed by our response.

Anonymous Referee #2

R2.1. Introduction and Conclusions and summary: Please better highlight the advancements respect to previous work by Gangrade et al. (2019), Journal of Hydrology https://doi.org/10.1016/j.jhydrol.2019.06.027

Our response:

The main difference between Gangrade et al. (2019) and this study is on the type of flood events. Gangrade et al. (2019) focused on evaluating flood risks associated with probable maximum flood events (AEP < 10^{-4}) that are considered as the physics-based upper bound of surface inundation. In this study, we focused on more frequent extreme streamflow events (i.e., AEP around 1–0.2%) that are involved in broader engineering applications. Given their distinct nature and AEPs, different analyses and modeling strategies are hence needed. Also, the study area is different between these two studies. We have clarified the differences between Gangrade et al. (2019) and current study in various locations of this revised manuscript.

R2.2. L 365 referring to Fig. 2: the control and baseline samples of annual maximum peak streamflow (box-plots) may be seen as "significantly different" rather than "comparable". Indeed, two points need to be clarified in this respect: a) the shown baseline sample is relative to bias-corrected data or not? b) control and baseline samples have different lengths, so, perhaps a more objective way of comparing them may be to apply some bootstrapping algorithm, or to randomly extract from the baseline sample several sub-samples having the same length of the control sample, and compare these somehow.

Our response:

Thank you for the insightful and constructive suggestion. In addition to the box plot comparison, we have also conducted a two-sample t-test ($\alpha = 0.05$) to compare if the means of control and baseline annual maximum streamflow are statistically significantly different. The two-tailed analysis resulted in a p-value of 0.093 (see Table R1), which suggests that there is no significant difference between the means and support the statement that the control and baseline annual maximum streamflow values are comparable. To address this concern, the following statement has been included in the manuscript (Line 374): "*In addition, we also conducted a two-tailed two-sample t-test ($\alpha = 0.05$) to compare if the means of control and baseline annual maximum* *streamflow are statistically different. The results yielded a p-value of 0.09 which suggested that there is no significant difference between the means of both control and baseline simulations.*"

Table R1 - t-Test: Two-Sample Assuming Unequal Variances

	Control	*Baseline*
Mean	331.92	388.23
Variance	30955.45	26239.07
Observations (n)*	32	440
Hypothesized Mean Difference	0	
df	34	
t Stat	-1.73	
t Critical two-tail	2.032	
P two-tail	0.093	

* (n-1) degrees of freedom used for standard deviation calculation

R2.3. Fig. 8. It may be possible to derive the analogous curve for the control scenario hydrographs. How does this compare to the shown baseline and future curves?

Our response:

Although numerically speaking one may derive a similar curve only based on 32 years of control simulation results, it can be misleading and biased given the much smaller sample size. The baseline and future curves in Fig. 8 were each derived from 440 years of data to reduce uncertainty and render more confidence in the evaluation of lower AEP estimates. However, the control scenario consists of only 32 years of data which may not be suitable to support a meaningful comparison in Fig. 8.

R2.4. L 188: Many researchers consider as a standard choice a period of 30 years instead of 40 years. A comment on this may be added to the MS.

Our response:

Based on our understanding, a minimum of 30-year period was used in many studies (e.g., Alfieri et al., 2015a, 2015b) so that one may have a sufficiently long temporal window to capture the multi-decadal climate variability. Given the additional data provided by Gangrade et al. (2020), we have adopted a longer 40-year period which may further enlarge the sample space to better support the statistical analyses in this study. This additional clarification has been included in the revised manuscript.

R2.5. L 476: there is only an indirect demonstration that the model can reproduce well flow velocity. As no direct comparison is performed (data are not available in this sense, as far as I have understood), perhaps this should be downplayed.

Our response:

Thank you for the suggestion. The authors agree that the accuracy of simulated velocities is not evaluated within the scope of this study. To avoid overstatement, the manuscript has been modified by omitting "velocities" in the original statement. The statement now reads *"…demonstrating that the TRITON could reasonably estimate flood inundation extent and depths in the CRW.*"

R2.6. L298: A minimum threshold of 10 cm flood depth was used to judge whether a cell was dry or wet. How much do you think your results can be sensitive respect to this threshold value?

Our response:

Thank you for your comment. With regards to the threshold value, Gangrade et al. (2019) tested the sensitivity of the threshold value in their study. Aside from the minimum threshold of 10cm, they also tested a minimum threshold of 1cm and reported minimal impact in maximum inundation area. Based on this fact, we can reasonably conclude that the results are marginally sensitive to the threshold values lower than 10cm. It is also important to note that this threshold value only applies to the analysis of post-processing results, as the accuracy with which TRITON decides whether a cell is dry or wet during the calculation is 10^{-12} meters.

References:

Gangrade, S., Kao, S. C., Dullo, T. T., Kalyanapu, A. J., & Preston, B. L. (2019). Ensemble-based flood vulnerability assessment for probable maximum flood in a changing environment. *Journal of Hydrology*, *576*, 342–355. https://doi.org/10.1016/j.jhydrol.2019.06.027

Gangrade, S., Kao, S. C., & McManamay, R. A. (2020). Multi-model Hydroclimate Projections for the Alabama-Coosa-Tallapoosa River Basin in the Southeastern United States. *Scientific Reports*, *10*(1), 2870. https://doi.org/10.1038/s41598-020-59806-6