Issues remain between the study's aims, methods, and conclusions. This type of dataset is not appropriate to "map the spatial variability" (line 100 -101). As stated in previous comments, the gaging network does not provide a representative sample to map the spatial variability. The authors have added some discussion acknowledging some of these limitations. However, it does not remedy the disconnect between the aim and methodology. For instance, smaller streams are vastly under-represented; and some individual larger rivers contain larger portions of gages than those surrounding them. These rivers and streams comprise a population of geomorphic characteristics, that if biased in representation, impacts the results of spatial variability in geomorphic response. While mapping the predicted residuals at streams is valuable and inferences can be made by comparing results across different regions, the authors should reconsider the information provided by the methods relative to the study's aims. The authors fail to mention spatial patterns in the conclusions, except perhaps, a slight alluding to at line 695.

The introduction should be more specific about the type of response considered in this analysis: abrupt loss of channel capacity following flood inducing storm events. This will provide context and clarity.

Minor Comments:

- Line 57 remove parenthesis around citation.
- Line 131. Revise sentence by removing "and" and capitalizing first word in sentence.
- Line 197: Greater than "the" 80th percentile. And 80th percentile of what? The annual flow duration curve?
- Line 372: To me, the relative magnitude of the RMSE is unclear. The authors state that it was close to 0, which is good. But the values range from 0.09 – 0.14 m. What would be considered unacceptable? While zero is the target; I do not have a frame of reference to know the relative magnitude of this error. This is important, if I understand correctly, because it influences confidence in the predicted residuals. What confidence can we have in the predicted residuals?
- Line 605: Remove the text "In the revised manuscript, we will incorporate these comments."
- Line 609: I am unsure about translating the model to gaged sites. In my previous comment, I was referring to a limitation in data availability. Wouldn't the model require a stage-discharge curve to relate the residual to? I a stage-discharge curve would not be determined from gage extrapolation.
- Line 619: Was the ML model trained using geomorphic properties as well?