

Interactive comment on “Predicting probabilities of streamflow intermittency across a temperate mesoscale catchment” by Nils H. Kaplan et al.

Anonymous Referee #2

Received and published: 13 July 2020

This study presets a method to estimate streamflow intermittency across the temperate landscape. The method builds upon previous work adding a probability estimation. The end goal is to get towards spatially explicit mapping of streams to support representation in modeling efforts. The study focuses on the well-investigated Attert Catchment in Luxembourg. The topic is timely given potential extremes and change brought about by climate shifts in landscape-scale hydrological function. The paper is well written and easy to follow. With that, there are only a few comments that need to be addressed in a revision. Addressing these will take some work, but nothing too laborious and should strengthen the study.

The first aspect that should be addressed would be the relative coarse resolution of the DEM of 15 m resolution impact on model uncertainty and/or sensitivity. The authors

C1

point out the potential shortcoming due to the resolution and highlight that no higher resolution is available. It still seems more is needed here to support the potential role of the coarse resolution. For example, missing even the smallest of surface channel flow to connect sections of flow streams is problematic for the modeling approach. So, of course, I'm left wondering on the impact for the modeling and estimation. Could there be some quantification of the potential uncertainty? For example, a pseudo/synthetic reduction (or increase) in coarseness of the resolution and a simple re-run to assess the change in accuracy? That could at least partly quantify impacts and try to put a number on it. I'm sure there are plenty of more creative alternatives, but putting some uncertainty bound (confidence) around your estimates as a function of the spatial resolution would be helpful since you are thinking to use your model output in connection with a modeling effort. Without tracking through the added uncertainty, you're likely to see some huge multiplier effects for follow up estimations.

The other aspect that jumps at me would be the lack of some validation using a leave one out or a split sampling to get at model sensitivity and robustness. There is much done to assess the model performance and consider the power of each separate model. But, could you provide some sort of validation of the accuracy? Seems a systematic leave-one-out-at-a-time approach could be useful to model a real data point and see if you got it right or wrong (and track the false positives to see if you are getting wet or dry too much). Alternative could be some Monte Carlo split samples to estimate several points left out of a training dataset - then randomize and repeat. Yes, these approaches are brute force and cumbersome, but this study is all desktop and computer based. So should be “simple” enough to add some loops and let the program churn out some validation statistics. That would help the reader assess how much the configuration of sample locations drives the accuracy and performance. Could you even assess the “value” of each observation point in the overall system to help future studies design where to sample intermittency for the most bang for the buck.

