Hydrol. Earth Syst. Sci. Discuss.,
https://doi.org/10.5194/hess-2020-181-AC2, 2020

Hydrology and Earth System Sciences
Discussions

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

Nils H. Kaplan et al.

nils.kaplan@hydrology.uni-freiburg.de

Received and published: 16 August 2020

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 be 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.»

Dear Referee,

Thank you for your helpful comments and suggestions to improve our manuscript. We appreciate your general agreement with the structure of the paper and content of our manuscript. Please find your questions and comments marked as e.g. « R2.C1: question/comment» followed by our answer marked as e.g. R2.A1: below. Best regards, Kaplan et al.

«R2.C1. Specific comments: The first aspect that should be address would be the relative coarse resolution of the DEM of 15 m resolution impact on model uncertainty and/or sensitivity. The authors point out the potential shortcoming due to the resolution and highlight the 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 »

R2.A1.

Good point. It was indeed unfortunate that no higher resolution DEM was available! However, the DEM used in this study is suitable for most of the sites also after careful shifting of some sites to the calculated stream channel. Nevertheless, 8 sites may be prone to non-accurate delineation of the catchment area, mainly in flat and/or areas with highly detailed relief and small drainage channels. We will make an effort to better assess the uncertainty caused by the coarse-resolution DEM and will include this in the revised manuscript.

«R1.C2. Specific comments: The other aspect the jumps at me would be the lack of some validation using a leave on 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 of wrong (and track the false positives to see if you are getting wet or dry too much). Alterative 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 be 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. »

R2.A2.

We will add the Leave One Out Cross Validation (LOOCV) approach and update the Methods section for the map comparison. Therefor we will change the following parts in the manuscript:

Methods: page 15 line 4:

[. . .] Due to the small data set which does not allow for a split validation approach, a leave on out cross validation (LOOCV) approach (e.g. Akbar et al., 2019; Ossa-Moreno

et al., 2019) was applied to validate the model based on the original data set. Therefore, for each observation n in the dataset one model is run with one observation left out from the dataset. Thereafter, the GLM derived from n-1 data points is used to predict the value y_hat for the left-out point y. This process is repeated for all observations. The measure of Root Mean

Square Error RMSE is used to assess the model accuracy as follows:

RMSE= $\sqrt{(1/n \sum_{(i=1)}^{n} (y_i - (\hat{y}_i))^2)}$

and bias of the model by:

Bias=1/n $\sum_{(i=1)}^{n} (y_i - (\hat{y}_i))$

where n is the number of observations, y_hat is the predicted and y the observed relative intermittency (Akbar et al., 2019).

Methods: page 15 line 8:

In order to validate the results from the classified reaches we compare the stream length of the modelled streams with the length of the streams from the topographic map (Le Gouverment du Grand-Duché de Luxembourg, 2009). We assume that the mapped stream network approximately represents the natural layout of the stream network in areas with lower human impacts.

On page 17, line 4 (Results section):

The RMSE for the Model-Y is with 0.26 the lowest among all models with RMSE of 0.263 and 0.29 for the Models W2 and D2 respectively. The bias of the models is very low and ranges around zero with values between - 9.6 * 10-4 (Model Y) and 4 * 10-5 (Model W2). Model deviations for the models Y, D2 and W2 is shown in fig. 7.

On page 23 line 33 (discussion section): Instead of:

"As the distribution of measurement sites in the data set has a strong tendency towards permanent streamflow sites and thus to the perennial reaches, this leads to an underrepresentation of the intermittent and ephemeral reaches in the data when splitting the data for model evaluation (Figure 6). We therefore evaluated the model by its ability to predict the spatial distribution of intermittent/perennial streams compared to the mapped stream network. We assume that the mapped stream network approximately represents the natural layout of the stream network in areas with lower human impacts. However, alteration of the natural stream network in areas of artificial and agricultural land use can be severe and thus misleading when comparing to model results."

we will write:

"[. . .] good GLM setup. However, the leave- one-out cross validation allows for a data-based validation. The RMSE values (0.26 – 0.31) obtained for the different models related to the maximum possible RMSE of 1 shows overall model deviations of around 26 to 30 %. The plotted deviations (fig. 7) reveal some extreme deviations of nearly 1. The majority of perennial streams seem to be well represented by the model, while many of the ephemeral streams have deviations of > 0.5. This could be due to the distribution of observation sites in the data set, which have a strong tendency towards permanent streamflow sites and thus to the perennial reaches, while intermittent and ephemeral reaches are underrepresented (Figure 6). The better representation of perennial streams becomes also visible in the model validation by its ability to predict the spatial distribution of intermittent/perennial streams compared to the mapped stream network."

Fig. 7 Full caption:

Figure 7: Model deviations for all models. The indicated intermittency class is based on the classification scheme from Hedman & Osterkamp (1982). The observed relative intermittency is < 0.1 for the ephemeral, $\geq$0.1 and < 0.8 for the intermittent and $\geq$ 0.8 for the perennial class.

Reference update:

Akbar S., Kathuria A. and Maheshwari B.: Combining imaging techniques with non-parametric modelling to predict seepage hotspots in irrigation channels. Irrigation Science, 37, 11–23, doi: 10.1007/s00271-018-0596-6, 2019.

Ossa-Moreno J., Keir G., McIntyre N., Cameletti M. and Rivera D.: Comparison of approaches to interpolating climate observations in steep terrain with low-density gauging networks, Hydrol. Earth Syst. Sci., 23, 4763–4781 doi: 10.5194/hess-23-4763-2019, 2019.

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., https://doi.org/10.5194/hess-2020-181, 2020.

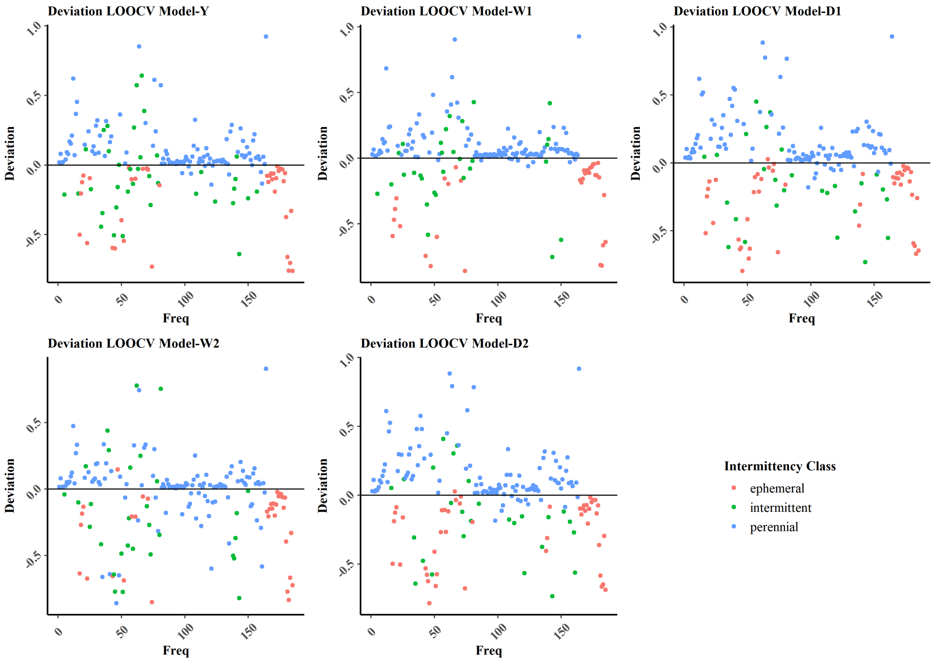

Fig. 1. Figure 7: Model deviations for all models. The indicated intermittency class is based on the classification scheme from Hedman & Osterkamp (1982). The observed relative intermittency is < 0.1 for the