Hydrol. Earth Syst. Sci. Discuss.,
https://doi.org/10.5194/hess-2019-650-AC2, 2020

Hydrology and Earth System Sciences Discussions
Open Access

Interactive comment on "Predicting tile drainage discharge using machine learning algorithms" *by* Saghar Khodadad Motarjemi et al.

Saghar Khodadad Motarjemi et al.

sa.m@agro.au.dk

Received and published: 21 February 2020

Dear Dr. Stamm,

Thank you so much for the remarks and below you can find Authors' reply to each comment.

1: Editor comment:

L. 83: The year of the citation of Wong is inconsistent between the main text and the reference list.

Authors' reply:

The correct citation year is 2015; however, the reference will be removed from the

manuscript, as it was misplaced in line 83.

2: Editor comment:

L. 156 - 157 The number of models seems to be inconsistent.

Authors' reply:

In total, two algorithms were used and six models were trained with three different resampling or cross-validation methods. The correct number is six and line 157 will be corrected.

3: Editor comment:

L. 190 - 191: What could be reasons why KF performs worse? Is there a spatial bias?

Authors' reply:

As indicated on Lines 226, 265 and presented on Table 2, the highest accuracy was achieved with k-fold (KF) cross-validated models. The lines 190 and 191 will be reformulated to better stress that KF yields a higher accuracy.

4: Editor comment: L. 238: What is the basis for this statement?

Authors' reply: Adhikari et al., 2013. The citation is missing in this line and will be added to the revised manuscript.

5: Editor comment:

L. 244: Does the predictive value of DEM-derived indices not depend very much on the spatial support and resolution of the data? Have you calculated these indices as averages across the catchments?

Authors' reply:

In the manuscript, we used the DEM-derived covariates at the point of the drainage outlet. We agree that it would be more useful to use averages for the drainage catchments, however, information on the extents of the drainage catchments was not consistently available, and we would have had to exclude a large number of stations to use this approach. The reasoning will be also stated on the revised manuscript.

6: Editor comment:

Fig. 1a: The data seem to separate into two clusters. Do the points with high discharge but rather low percolation have something in common that could explain the differences?

Authors' reply:

The figure, which contains the information, is Fig. 2a. Due to the different catchment sizes, the discharge behavior might differ between large and small catchment. For the larger catchments, discharge generated in the pipes might not necessarily flow to the outlet but might re-infiltrate into the soil depending on the spatial variability of the soil in the catchment (e.g. areas that are sandier where the natural drainage capacity (drainage class) of the soil is higher). Some of the drainage stations are draining large catchments, which could explain the clustering (Fig. 2b) when the percolation (Db) is compared to drainage discharge (Q).

7: Editor comment:

Fig. 1a&b: Combining the two data suggests that drainage discharge is well correlated (and predicted) by the amount of precipitation. How does this relationship look like if you additionally distinguish between clay and sandy soils?

Authors' reply:

The Figure, which contains the information, is Fig. 2a&b. Fig. 2a demonstrates the correlation between measured drainage discharge (Q) and calculated percolation (Db), and Fig. 2b shows the correlation between measured precipitation (P) and calculated Db. Here we attach an extra plot where the correlation between measured P and Q is demonstrated (Figure 1). We have included the soil type as a predictor in the model

but an extra figure showing the relation between discharge and clay percent could be included in the revised manuscript. The analysis to distinguish between clay and sandy soils will be carried out and included in the revised manuscript.

8: Editor comment:

Tab. 1: Please provide the distribution of predictors (as supporting information).

Authors' reply:

Mean values for all the covariates, excluding the categorical ones, is inserted in Table 1. Based on the comments from Referee #1, depth of sinks (BS) will be excluded from the covariates.

Kind Regards,

Authors

Please also note the supplement to this comment:
https://www.hydrol-earth-syst-sci-discuss.net/hess-2019-650/hess-2019-650-AC2-supplement.pdf

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., https://doi.org/10.5194/hess-2019-650, 2020.

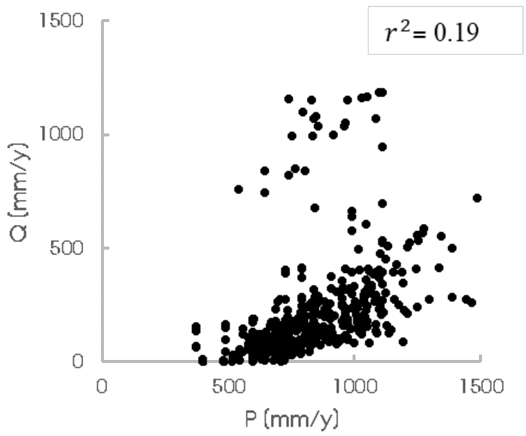

Fig. 1. Correlation between measured precipitation (P) and measured drainage discharge (Q)

Predictors	Description	Range/ Class	Mean
Db	Percolation/Discharge out of the root zone (mm y^{-1})	0 – 1033	336
Geological_R	Geological region	7 classes	-
DEM	Elevation (m)	0.74 – 83.16	30.53
Geological_C	Geology of the area	10 classes	-
F_Accu	Flow Accumulation/Number of unslope cells	1 – 1108	14.48
SagaWI	SAGA Wetness Index	12.16 - 16.58	14.17
TWI	Topographic Wetness Index	3.47 – 12.33	5.43
D_Class	Drainage class	5 classes	-
Clay A %†	Clay content 0-30 cm soil depth	3 – 20.3	13.29
Clay B %†	Clay content 30-60 cm soil depth	2 – 29.1	18.22
Clay C %†	Clay content 60-100 cm soil depth	1.5 – 31	19.5
Clay D %†	Clay content 100-200 cm soil depth	2.2 – 32.6	18.91
DDJD LER-A%‡	Clay content in A horizon	3 – 24.8	14.34
DDJD LER-B%‡	Clay content in B horizon	0 – 31.97	18.46
DDJD LER-C%‡	Clay content in C horizon	0 – 29.1	20.94
JB	Danish soil classification for the A horizon	12 classes	-
Gwd_Int	Depth to groundwater table interpolated from well observations and surface water (m)	0 – 25.31	7.42
Wetlands	0: Non-wetlands; 1: Wetlands; 2: Central wetlands; 3: Peatlands.	4 classes	-
D_DK_New	Artifical drainage-new map	2 classes	
DP_New	Drainage probability-new map	0 – 0.86	0.72
D_DK	Artifcial drainage-old map	2 classes	-
DP	Drainage probability-old map	0 – 0.82	0.72
Demdetrend	Elevation minus the mean elevation in a 4 km radius (m)	-11.4 – 26.04	6.23
Dirinsola	Direct insolation (kWh/year)	1150.08 – 1348.61	1273
Gwd_model	Depth to groundwater from the model (m)	0 – 32.42	5.54
Hdtochn	Horizontal distance to the nearest waterbody (m)	0 – 1114.89	324
Midslppos	Mid-slope position	0 – 0.7	0.25
Mrvbf	Multi-resolution index of valley bottom flatness	0.07 – 8.68	3.69
Slpdeg	Surface slope gradient (degrees)	0.09 – 7.53	1.46
Slptochn	Downhill gradient to the nearest waterbody (m)	0 – 3.48	1.20
Vdtochn	Vertical distance to the nearest waterbody (m)	0 – 19.28	6.11
Valldepth	Valley depth (m)	2.43 – 21.35	4.97
Landscape	Landform types	11 classes	-

Fig. 2. Table 1. List of covariates used to predict the discharge including a description of the parameter and a range specifying the type of covariate.