We tried to do the predictions with fewer and also with more than 5 variables. In some cases, the error lowered with the change in predictor variable number, but in some cases, the error got higher. Unfortunately, no connection could be found between the OOB error and the number of variables. The results are presented through the example of phytobenthos in Fig. 1.

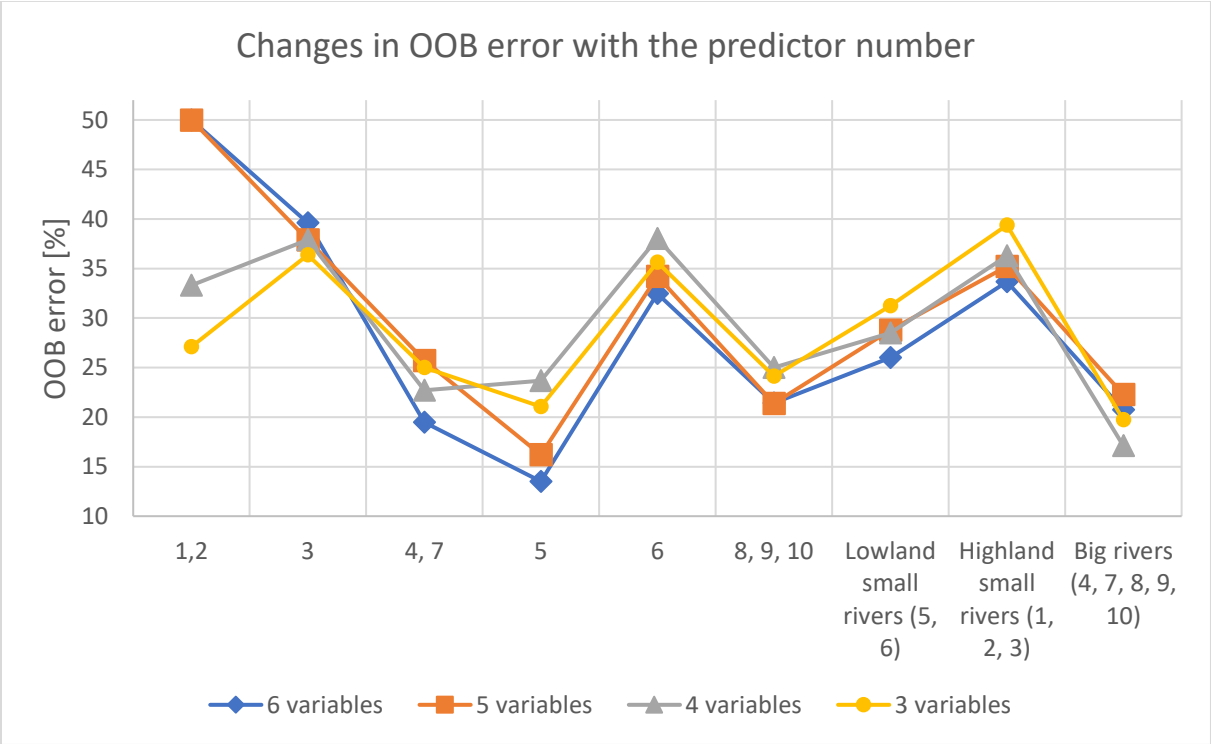

Figure 1. Changes in OOB error with the predictor variable number in the case of binary classifications of phytobenthos.

As can be seen in the figure, the size of the error is not proportional to the number of variables, and an optimal function cannot be determined in general. Even when the figures with the variable importance indicate that the right number is not 5, the error might be bigger or smaller. For example, in the case of highland small rivers (types 1, 2, 3) according to the variable importance plot (Fig. 4. in the manuscript) BOD_5 concentration, dissolved oxygen concentration and proportion of extensively used pastures are the most important ones standing apart from the rest, and all the other variables have quasi the same importance. Despite this, OOB error gets higher when only the top 3 variables are considered.

Type 1, 2 reacted differently to the change in the number of variables, as this type also had the largest prediction error.

Although no clear rule can be established, it can be seen that the best prediction was achieved with 4-6 variables with one exception. This justifies the arbitrary choice of 5 variables.

The variable ranking used for assessing the changes in OOB error can be seen in Table 1.

SUPPLEMENTARY MATERIAL S2

	River type								
Variable ranking	**1, 2**	**3**	**4, 7**	**5**	**6**	**8, 9, 10**	**Lowland small rivers (5, 6)**	**Highland small rivers (1, 2, 3)**	**Big rivers (4, 7, 8, 9, 10)**
1.	Dissolved oxygen	BOD_5	Chloride	NO_2-N	Ext. used pastures %	NO_3-N	Electrical conductivity	BOD_5	Continuity status class
2.	Water temperature	Ext. used pastures %	Orthophosphate	Orthophosphate	Electrical conductivity	Total N	Chloride	Dissolved oxygen	Suspended matter
3.	BOD_5	Dissolved oxygen	Total P	Total P	NH_4-N	Inorganic N	Orthophosphate	Ext. used pastures %	Total P
4.	Organic N	NO_2-N	Suspended matter	Ext. used pastures %	Water temperature	TOC	BOD_5	Electrical conductivity	Orthophosphate
5.	Chloride	NH_4-N	TOC	BOD_5	Orthophosphate	Orthophosphate	NO_2-N	Organic N	TOC
6.	COD_{Mn}	Water temperature	Continuity status class	Dissolved oxygen	BOD_5	Dissolved oxygen	Ext. used pastures %	Intensive agricultural areas	BOD_5

Table 1. Variable ranking used for assessing the changes in OOB error with the predictor variable number in the case of binary classifications of phytobenthos.