

Hydrol. Earth Syst. Sci. Discuss., author comment AC2
<https://doi.org/10.5194/hess-2022-141-AC2>, 2022
© Author(s) 2022. This work is distributed under
the Creative Commons Attribution 4.0 License.

Reply on RC2

Johannes Laimighofer et al.

Author comment on "Low-flow estimation beyond the mean – expectile loss and extreme gradient boosting for spatiotemporal low-flow prediction in Austria" by Johannes Laimighofer et al., Hydrol. Earth Syst. Sci. Discuss.,
<https://doi.org/10.5194/hess-2022-141-AC2>, 2022

Dear Reviewer#2

We want to thank the reviewer for his positive feedback.

We will address all points raised in the review in the final author response. Here we give just a short remark on your point about potential collinearity.

"My only concern is the potential existence of collinearity of temporal predictors as the authors considered CWB, CWB_center, and CWB_SDI all together as potential predictors for finding the best model. It would be good to see an analysis about if including collinear predictors yields a significant increase in the model performance compared to when only not collinear predictors are considered for the model fitting."

Thank you for this comment. Concerning our preselection of temporal predictors, we were aware that not all temporal variables would be necessary in the sense that they will improve the predictive performance of the model, in the light of collinearity. However, the XGBoost model (Chen et al. 2016), and boosting in general (Friedmann 2001, Hastie et al. 2009, especially when used in a nested CV-approach) is known to handle collinearity of predictors through regularization parameters in a highly sophisticated way. In our approach, the inner CV loop further assures that predictors are only selected if they increase the predictive performance of the model, and the outer loop, additionally, evaluates the predictive performance at ungauged sites independently from model fitting. We have further checked that only using the CWB and the different lags yield similar performance to the presented model. For these reasons we can safely argue that collinearity is explicitly handled in our approach, and we will clarify this point in the revised MS.

References

Chen, T. and Guestrin, C.: XGBoost: A Scalable Tree Boosting System, KDD '16, p. 785–794, Association for Computing Machinery, New York, NY, USA,
<https://doi.org/10.1145/2939672.2939785>, 2016

Friedman, Jerome H. "Greedy Function Approximation: A Gradient Boosting Machine." *The Annals of Statistics* 29, no. 5 (2001): 1189–1232. <http://www.jstor.org/stable/2699986>.

Hastie, T., Tibshirani, R., and Friedman, J. (Eds.): *The elements of statistical learning*, vol. 2, Springer series in statistics New York, Springer, New York, <https://doi.org/10.1007/978-0-387-84858-7>, 2009.