EGU23-4183
EGU General Assembly 2023

Time-series Boosting in Ensemble Modelling of Real-Time Flood Forecasting Application

Farshad Piadeh[1], Farzad Piadeh[2], and Kourosh Behzadian[2]
[1]Azad University of Mashhad, Computer engineering, Mashhad, Iran
[2]School of Computing and Engineering, University of West London, St Mary's Rd, London, W5 5RF, UK

While concept of boosting ensemble data mining techniques has been recently attracted a lot of attention for flood forecasting, mainly on non-urbanised river basins or reservoirs [1,2], time-series boosting, i.e., contribution of last timestep prediction to the next forecasting model is a new era, especially for real-time operation of flood forecasting models in the shape of early warning systems.

This study aims to provide time-series boosting for ensemble flood forecasting model through adding forecasted water level of one timestep before as an input of training base models to previous proposed rainfall feature, especially rainfall duration, intensity, evidence of past rainfall and season occurrence [3]. Several weak learner data mining techniques are developed for various forecast lead times and recorded in data cube structure that can be used for developing time-series boosted ensemble model. This novel model was tested for real case study of Hanwell urban drainage systems located in the west London, UK for a period of 20 years data with 15min intervals. Confusion matrix is employed for performance assessment and the model is compared by conventional benchmark gradient boosted models.

Results shows the added feature can significantly increase the accuracy of overflow detection of all developed base models, especially for longer timesteps. More specifically, adding the new feature to the model can increase the accuracy rate from 84% for the best developed base model to 93% in 3hrs-ahead predictions. More importantly, the model can decrease underestimation miss rate from 45% to only 21% for the same forecast lead time. Furthermore, new time-series boosted ensemble model can noticeably increase overflow detection rate, where hit rate increase from 78% to 88% in 3hrs-ahead predictions. Overall, the concept of time-series boosted ensemble modelling can overcome the problem of missing and false alarm of real-time operation by adding the previous situation of catchment to the forecasting procedure.

References

[1] Jarajapu, D., Rathinasamy, M., Agarwal, A., Bronstert, A. (2022). Design flood estimation using extreme Gradient Boosting-based on Bayesian optimization, *Journal of Hydrology*, 613(A), 128341.

[2] Piadeh, F., Behzadian, K., Alani, A. (2022). A critical review of real-time modelling of flood

forecasting in urban drainage systems. *Journal of Hydrology*, 607, 127476.

[3] Piadeh, F., Behzadian, K., Alani, A.M. (2022). Multi-Step Flood Forecasting in Urban Drainage Systems Using Time-series Data Mining Techniques. *Water Efficiency Conference*, West Indies, Trinidad and Tobago. repository.uwl.ac.uk/id/eprint/9690 [Accessed 31/12/2022].