Hydrol. Earth Syst. Sci. Discuss., 10, C132–C134, 2013
www.hydrol-earth-syst-sci-discuss.net/10/C132/2013/

Hydrology and Earth System Sciences Discussions

Interactive comment on "Assessing the predictive capability of randomized tree-based ensembles in streamflow modelling" *by* S. Galelli and A. Castelletti

R. Romanowicz (Referee)

romanowicz@igf.edu.pl

Received and published: 24 February 2013

The authors present a comparison of the predictive capability of extremely randomized trees (Extra-Trees), including the accuracy, computational efficiency and explanation ability, with other tree-based methods and data-driven approaches in modelling rainfall runoff processes.

The paper is of interest to the hydrological community. It is well written and well structured. Its main novelty lies in the analysis of the applicability of the Extra-Trees approach to stream flow modelling. The authors follow a four-step assessment procedure, including random sampling of observational data, and a multi-criteria evaluation

of model performance and they provide an uncertainty analysis of model performance. The method is applied to two catchments, a small urban Marina catchment in Singapore and the Cunning River, a large natural catchment in Australia. The authors compare the predictive capability, explanation ability and computational efficiency of the studied method to other data-driven methods (M5, CART, ANN and MLR). In particular, they compare the scatter plots of predicted and measured flows using all five methods and analyse the probability distributions of residuals.

The authors claim that the Extra-Trees method out-performs the other methods. Its largest advantage lies in its ability to rank the importance of the model input variables. Otherwise, its superiority is not so obvious. The M5 method has better results for all goodness-of-fit criteria. When it comes to computer efficiency, both Extra-Trees and M5 have much bigger computer time requirements than other methods. The authors are asked either to change their Conclusions or explain how they can justify their statement. A detailed examination of the scatter plots (Fig. 5) shows that Extra-Trees under-predicts high flows for the Marina catchment. However, it is not visible in the error distribution shown in Fig. 7. The authors show the results of the fitted logistic distribution, which hides the model performance. It would be more informative if the empirical distributions were also shown. It is also not clear what time periods were used during the comparison of the scatter plots and the error distributions. The reader assumes that they were the same, but this should be clearly stated.

In summary, the performance of the Extra-Trees method in predicting flow is the most interesting to the hydrological community, and that subject should be the main focus of the paper. At the moment, this particular point is not well explored, and the advantages of using the method are not convincing. In particular, the authors should comment on the ability of the methods they compare to reproduce the flow patterns, with a discussion of the model performance for high and low flows. The results obtained from the ranking of input variables look promising, but the outcome seems to be too obvious. For the Marina catchment, the influence of rainfall is the highest due to small catch-

ment retention. For the Cunning River, the flow values in two previous days have the largest impact on flow predictions. The results suggest that Extra-Trees could be useful in exploring the dependence of input ranking on flow under varying meteorological conditions. It would also be advantageous if the authors showed the predictions with the confidence limits for the Extra-Trees method. The information contained in the error analysis does not give a proper perspective.

An additional, important comment is on the quality of the figures. The figure labels are too small. I had to take a magnifying glass to be able to see the details.

The paper is suitable for publication in HESS after moderate corrections.

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., 10, 1617, 2013.