Investigating the Upper Limit of ML-Based Stochastic Discharge Simulations

Sonja Jankowfsky*[1], Ashish Kumar[1], Edom Moges[1], Shuangcai Li[1], and Arno Hilberts[1]

[1]Moody's – Moody's – France

Abstract

Machine learning (ML) hydrological models, particularly Long Short-Term Memory (LSTM) networks, have shown remarkable potential in accurately simulating streamflow. This study explores the capacity of regional ML models to extrapolate beyond their training data range by integrating stochastic precipitation from a weather generator into an LSTM hydrological model. Previous research by Jankowfsky et al. (2025) demonstrated that LSTM models extrapolate well with stochastic precipitation when a sufficiently large training dataset is used for gauges.

In this study, we aim to determine whether this finding holds for gauges near the upper limit of specific discharge (discharge per upstream area) of the modelled range. We utilize global and continental datasets comprising several thousand discharge gauges to conduct our investigation. The study examines the impact of different model structures, objective functions, and scaling techniques on the upper specific discharge limit. These tests reveal the importance of scaling techniques on the extrapolation capacity of the model. The study also confirms the earlier findings that the size of the training data set and the training strategy has a major influence on the upper limit of the model. Furthermore, as the specific discharge increases with smaller upstream area, it is important to include gauges on small tributaries into the training data set. Our tests show that with the right configuration, the upper limit can be pushed to reasonable physical limits allowing the model to be used for stochastic simulations.

Jankowfsky, S., Kanneganti, G., Li, S., Hilberts, A. Assteerawatt, A., Evaluation of LSTM Model for Stochastic Discharge Simulation. EGU General Assembly 2025, Vienna, Austria, 27 Apr–2 May 2025, https://doi.org/10.5194/egusphere-egu25-15152.

Keywords: Machine learning, hydrology, discharge simulation

*Speaker