EGU25-5863, updated on 21 May 2025
https://doi.org/10.5194/egusphere-egu25-5863
EGU General Assembly 2025

Long Short-Term Memory networks in hydrology: From free-time project to Google's operational flood forecasting model

Frederik Kratzert
Google, Research, Austria (kratzert@google.com)

Long Short-Term Memory networks (LSTMs) have been around since the early 90's but only in the last few years have LSTMs gained significant popularity in the hydrological sciences. Related publication counts have grown exponentially, and LSTMs power some of the largest-scale operational flood forecasting systems.

In this presentation, I'll look back at my relatively short career as a student and researcher at the intersection of hydrology and machine learning. I don't claim to have introduced LSTMs to hydrology, but I'll share my own experience helping to develop this modeling approach into what it is today. We will look at what I saw in this neural network architecture, and why I thought it was well suited for hydrologic applications.

The tale goes as follows*: Once upon a time, in a land (not so) far far away, a (not so young) master student of environmental engineering was teaching himself the dark arts of machine learning (ML). While studying ML for automated fish detection, he stumbled upon the LSTM architecture. Having just concluded a course on the design of conceptual hydrological models, he noticed the underlying similarity between the LSTM and these established approaches — and more generally, the conceptual approach for modeling the water cycle. With one of his dearest colleagues and friends, he started to work night and day (actually more nights than days) to see if the LSTM is indeed suitable for hydrology. From initial attempts at emulating the ABC and HBV models, to first real-world experiments in individual catchments, the LSTM was showing great potential. But it was not until he discovered the CAMELS dataset and started experimenting with large-sample hydrology that he fully understood the potential of LSTMs for applications in hydrology. Equipped with nothing more than his first GPU, he embarked on a quest to explore the wondrous lands of academia. Countless nights were spent on the computer, forging transatlantic friendships, conducting experiments and writing publications. Eventually, he ascended to the ranks of PhDs by defending his research against Reviewer #2 and the high council of the PhD committee. Fast forward in time, today, LSTMs are widely used and among others, power Google's current operational, global-scale flood forecasting model. And thus, the now (not so) old research scientist lived happily ever after with his wife and his children, and continues, to this day, to do much the same as he had in those earlier years.*

If there is one thing that I would like for you to take away from this talk, it is that I hope my

presentation will motivate young scientists to stay curious, to follow their own ideas, to not get demotivated by initial pushback and to not be afraid of reaching out to more senior researchers. I want to advocate strongly the importance of open science, of reproducibility, of collaborations, of benchmarking and of open data sharing to advance science.