Rainfall-Runoff Simulation and Interpretation in Great Britain using LSTMs

Thomas Lees¹, Marcus Buechel¹, Bailey Anderson¹, Louise Slater¹, Steven Reece³, Gemma Coxon⁴, and Simon Dadson¹,²
¹Oxford University, School of Geography and Environment, Oxford, United Kingdom of Great Britain – England, Scotland, Wales (thomas.lees@chch.ox.ac.uk)
²NERC Centre for Ecology and Hydrology, Wallingford, United Kingdom
³Department of Engineering, University of Oxford, Oxford, United Kingdom
⁴Geographical Sciences, University of Bristol, Bristol, United Kingdom

Techniques from the field of machine learning have shown considerable promise in rainfall-runoff modelling. This research offers three novel contributions to the advancement of this field: a study of the performance of LSTM based models in a GB hydrological context; a diagnosis of hydrological processes that data-driven models simulate well but conceptual models struggle with; and finally an exploration of methods for interpreting the internal cell states of the LSTMs.

In this study we train two deep learning models, a Long Short Term Memory (LSTM) Network and an Entity Aware LSTM (EALSTM), to simulate discharge for 518 catchments across Great Britain using a newly published dataset, CAMELS-GB. We demonstrate that the LSTM models are capable of simulating discharge for a large sample of catchments across Great Britain, achieving a mean catchment Nash-Sutcliffe Efficiency (NSE) of 0.88 for the LSTM and 0.86 for the EALSTM, where no stations have an NSE < 0. We compare these models against a series of conceptual models which have been externally calibrated and used as a benchmark (Lane et al., 2019).

Alongside robust performance for simulating discharge, we note the potential for data-driven methods to identify hydrological processes that are present in the underlying data, but the FUSE conceptual models are unable to capture. Therefore, we calculate the relative improvement of the LSTMs compared to the conceptual models, ΔNSE. We find that the largest improvement of the LSTM models compared to our benchmark is in the summer months and in the South East of Great Britain.

We also demonstrate that the internal “memory” of the LSTM correlates with soil moisture, despite the LSTM not receiving soil moisture as an input. This process of “concept-formation” offers three interesting findings. It provides a novel method for deriving soil moisture estimates. It suggests the LSTM is learning physically realistic representations of hydrological processes. Finally, this process of concept formation offers the potential to explore how the LSTM is able to produce accurate simulations of discharge, and the transformations that are learned from inputs (temperature, precipitation) to outputs (discharge).
References: