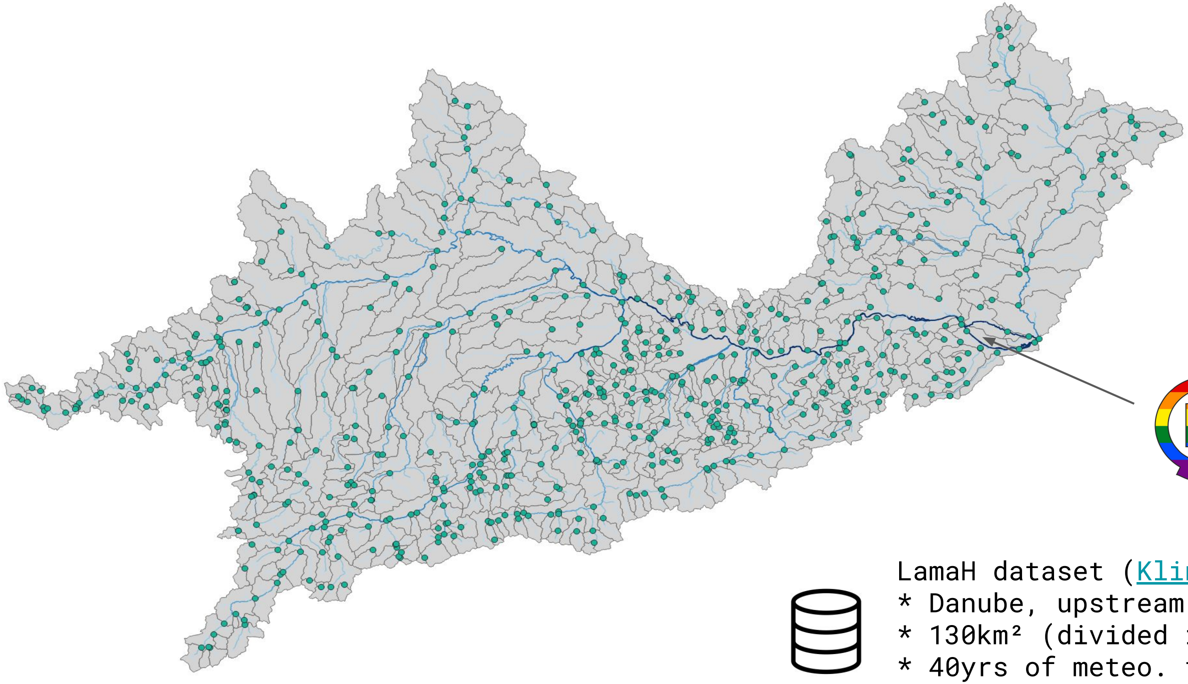
Large-scale river network modeling using Graph Neural Networks

Frederik Kratzert, D. Klotz, M. Gauch, C. Klingler, G. Nearing, and S. Hochreiter

@fkratzert E-mail: kratzert@ml.jku.at

Research question

Can we build a deep learning based, spatially distributed rainfall-runoff model to predict large river networks?



LamaH dataset ([Klingler et al., 2021 ESSDD](#))
* Danube, upstream of Bratislava
* 130km^2 (divided into 608 polygons)
* 40yrs of meteo. forcings and discharge obs
* static catchment attributes

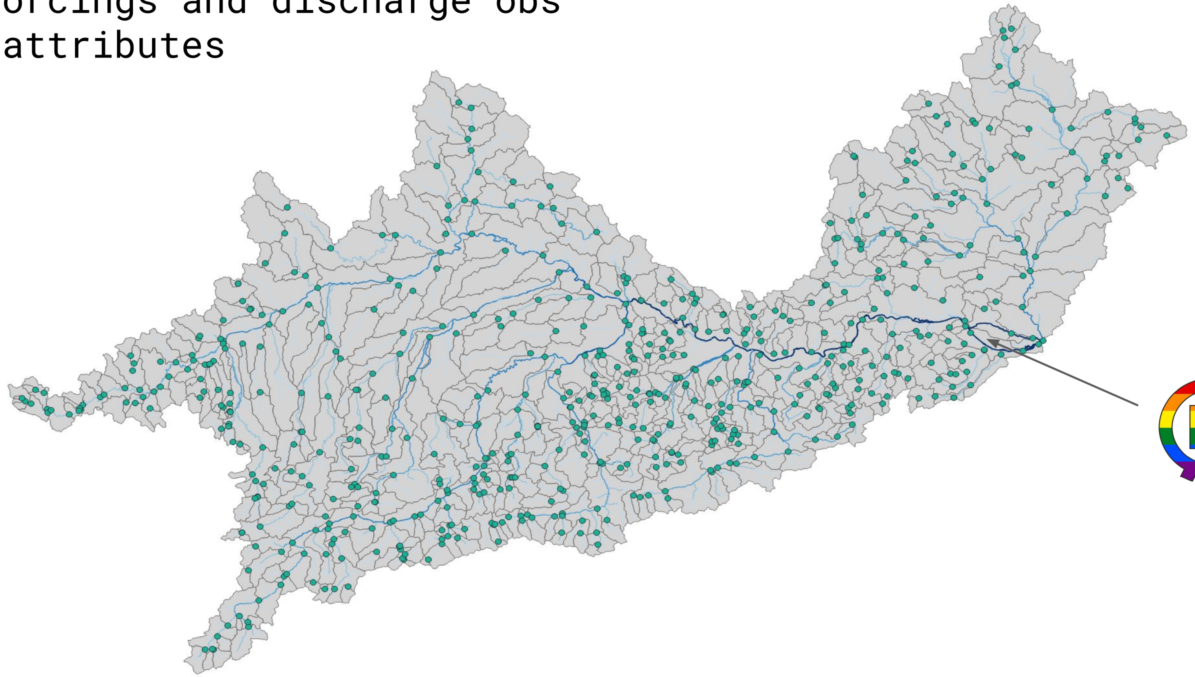
Research question

Why?

- Over the last years, deep learning models (especially the LSTM) have shown great potential in hydrology, excelling as rainfall-runoff models in different benchmark studies (e.g. REF1, REF2, REF3, REF4, REF5).
- One of the main benefits: deep learning models can learn anything that is deductible from data.
- In the context of a spatially-distributed rainfall-runoff models, this could be interesting to account for factors that are hard to include in conventional modelling approaches --- e.g. anthropogenic influences.

Data

LamaH dataset (Klingler et al., 2021 ESSDD)
* Danube, upstream of Bratislava
* 130km² (divided into 608 polygons)
* 40yrs of meteo. forcings and discharge obs
* static catchment attributes



Modeling concept part1: overview

Model consists of two parts:

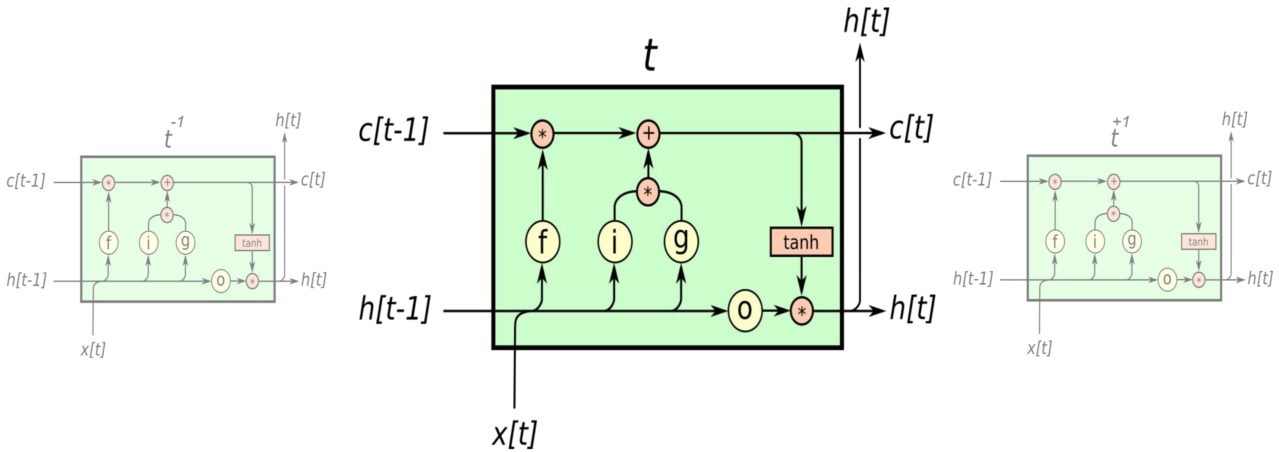
1. One model to predict the generated surface runoff of each polygon in the study area (polygon model).
2. Another model that learns to route the generated surface runoff along the river network (routing model).

Both are trained together, but it is also possible to initialize the polygon from a pre-trained basin model.

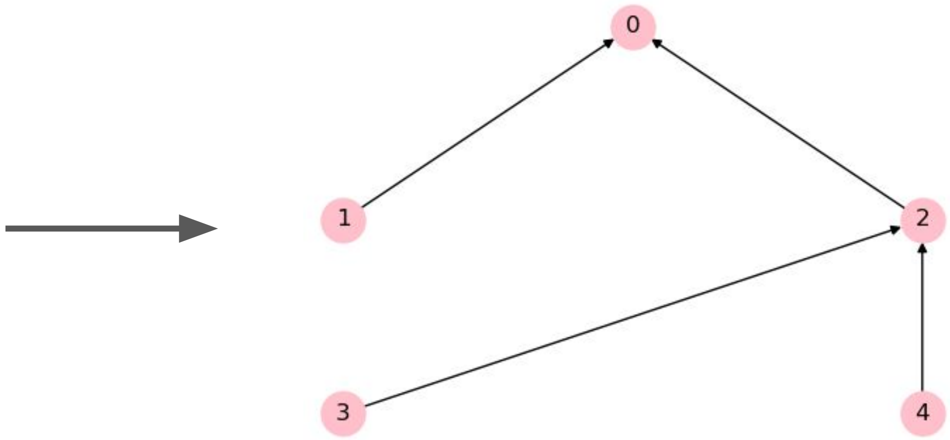
Modeling concept part 2: polygon model

*Surface runoff in each polygon is generated by a *single* LSTM, following the approach by [Kratzert et al. (2019)]:

*Inputs: meteor. forcings & catchment attributes
*Output: generated surface runoff



Graph Concept

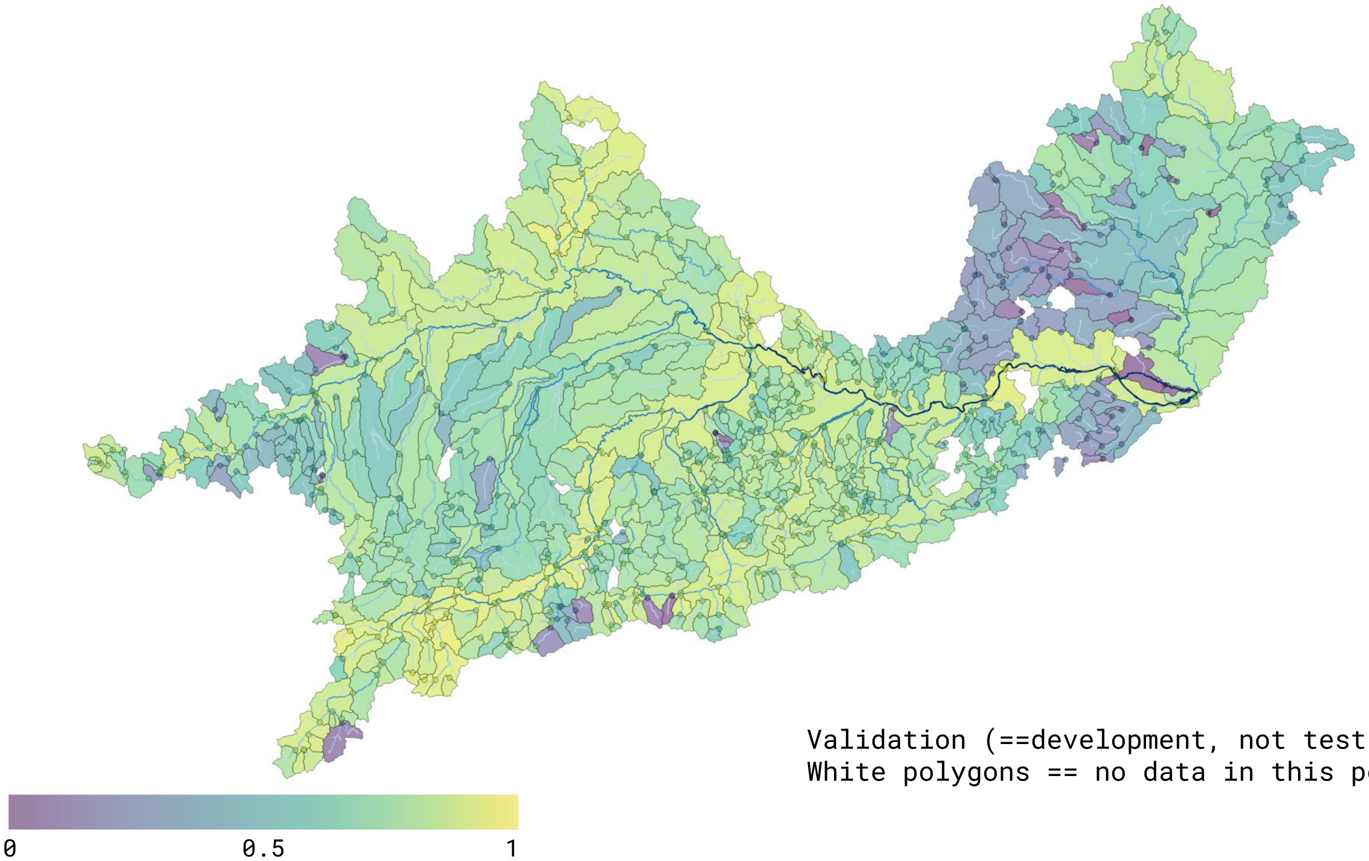


* every basin is one node in the graph.
* every river stretch between two polygon outlets (e.g. two gauges) is one edge in the graph.

Modeling concept part 3: routing model

- Generated surface runoff of the polygon model is used as node input.
- The routing is performed on the edge (i.e. the river stretch between two consecutive polygon outlets).
- Any additional inputs (static & dynamic) can be used.
- The actual routing model can be a linear reservoir or any differentiable function (e.g. any neural network).

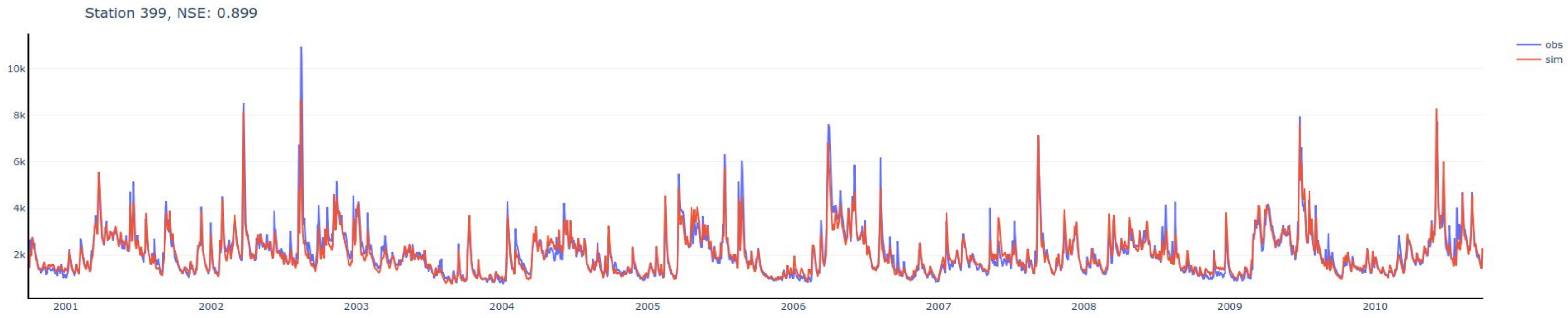
(very) preliminary results



Validation (==development, not test!) period results
White polygons == no data in this period

(very) preliminary results

Example of simulated discharge [m3/s] of one model for the most downstream gauge of the Danube (close to Bratislava).



Resources

Stay tuned for updates!

- Research blog: neuralhydrology.github.io
- Python library: neuralhydrology.readthedocs.io
- University institute homepage: jku.at/iml

Contact: kratzert@ml.jku.at