Deep learning for water quality prediction: the application of LSTM model to predict water quality in catchment scale

Amir Sahraei, Lutz Breuer, Philipp Kraft and Tobias Houska

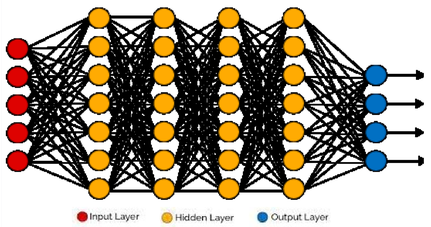
Institute of Landscape Ecology and Resource Management (ILR), Justus Liebig University Giessen, Germany

Motivation

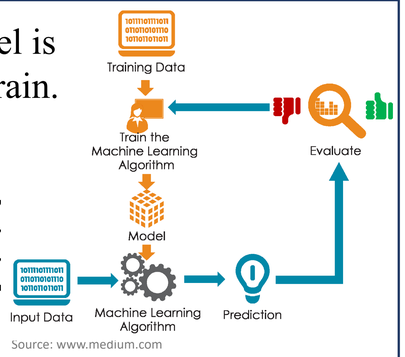
- The prediction of water quality is an efficient way for managing water resources and protecting ecosystems by providing an early warning against water quality deterioration.
- Process-based models are required extensive variety of parameters and are often limited by strict assumptions of boundary conditions and variable independence.
- The parameters are not easily obtained and it is extremely hard to produce adequate data at the catchment scale.
- A reliable and accurate estimate of water quality is not always ensured due to the such constraints.

Deep Learning

Neural network model is inspired by human brain.



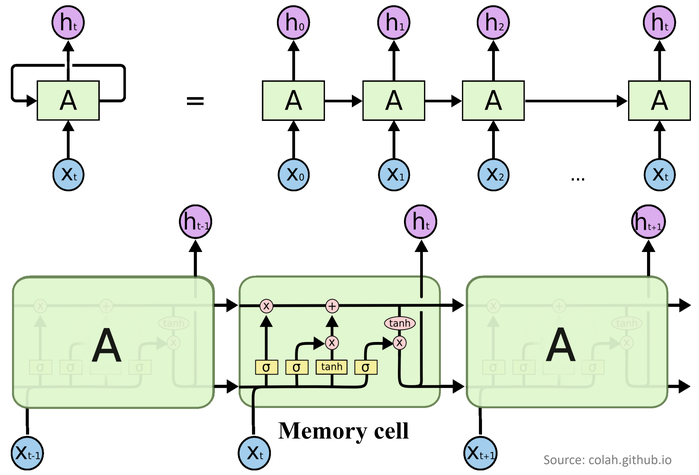
Deep Learning (DL) Neural Network



Scheme of modelling process

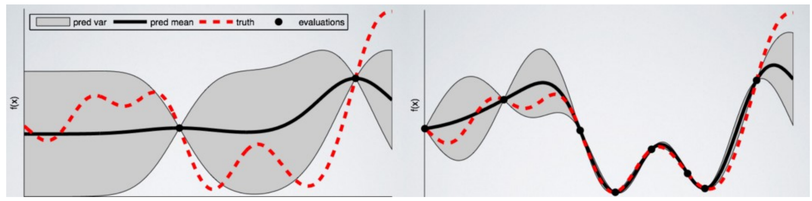
- DL is a class of machine learning algorithms with multiple layers that extracts high-level features.
- DL is able to efficiently predict complex and nonlinear system without any a priori knowledge about the underlying physically processes.

Long Short-Term Memory (LSTM)



Long Short-Term Memory schematic diagram

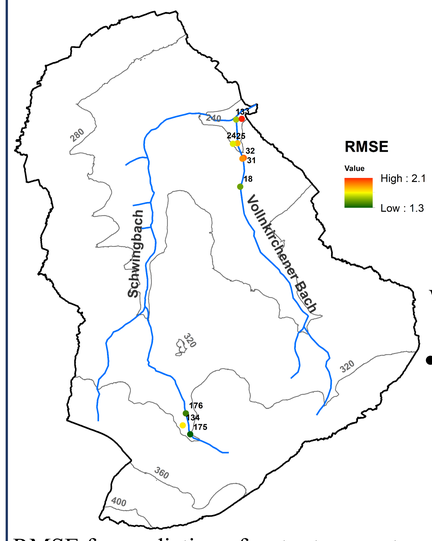
- LSTM network is a type of recurrent neural network (RNN) that is capable of learning long-term dependencies.
- **Memory cell** consists of forget, input and output gates that together control the flow of information within the LSTM network.
- **Forget gate** controls which information is removed from the cell state; **Input gate** defines which information is updated to the cell state; **Output gate** specifies which information is used from the cell state.
- **Bayesian Optimization** algorithm was used to optimize the model hyperparameters.



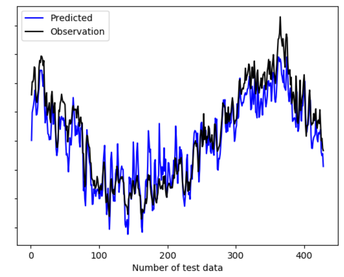
Results

Input variables: rainfall, air temperature, relative humidity and streamflow.

Target variables: N-NO_3, water temperature and water isotope concentration ($\delta^{18}O$) at 12 stream- and groundwater sampling sites over 2013-2020 historical data in Schwingbach Environmental Observatory (SEO), Germany.



RMSE for prediction of water temperature



Water temperature prediction at site #18

- A single LSTM model was trained over first six years to predict the water temperature of last year at the daily scale across the sites.

Conclusion and Outlook

- LSTM could successfully predict water quality parameters using only hydroclimatic data.
- Bayesian Hyperparameter Optimization is an efficient algorithm for the deep learning model development.
- Future research directions include applying the proposed model in other catchments with different characteristics (e.g., weather, topography and size).

amirhossein.sahraei@umwelt.uni-giessen.de