EGU21-8756
https://doi.org/10.5194/egusphere-egu21-8756
EGU General Assembly 2021

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
Justus Liebig Universität Gießen, Landscape, Water and Biogeochemical Cycles, Giessen, Germany
(amirhossein.sahraei@umwelt.uni-giessen.de)

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. So far, the classical approach is to predict water quality by the utilization of complex process-based water quality models. However, these models are not easy to set up and require comprehensive input data. The local characteristics, detailed process understandings and eventually data from land users such as farmers are needed, to build up a valid model structure. Such constraints can end up in wrong scientific conclusions ranging from false alarms to unpredicted environmental pollution in practical water monitoring application. Long short-term memory (LSTM) algorithms are known to be able to overcome some of the typical constraints in hydrological model applications. However, their performance in water quality prediction has rarely been explored. In this study, we investigate the ability of a LSTM model to predict the complex, nonlinear behavior of water quality parameters in the Schwingbach Environmental Observatory (SEO), Germany. We predict weekly nitrogen-nitrate concentrations, weekly stable isotopes of water concentrations ($\delta^{18}O$) and daily water temperature in six stream and six groundwater sources with different landuse and hillslope conditions. We use meteorological forcing data and catchment attributes as input variables. To ensure an efficient model performance, we employ a Bayesian optimization approach to optimize the hyperparameters of the LSTM. The model performance is evaluated by the Root Mean Squared Error (RMSE). Our LSTM is robust in capturing the dynamics of the water quality parameters over time. The RMSE for the LSTM performance ranges from 0.27 to 3.38 mg/l, from 0.069 to 0.27 ‰ and from 1.3 to 2.1 °C for nitrogen nitrate, $\delta^{18}O$ and water temperature, respectively. We compare the RMSE with statistical parameters of data. Results confirm that the LSTM is a promising tool for early risk assessment of water quality, particularly in view that only a minimal set of catchment information is needed to gain robust results.