EGU23-15604, updated on 02 Jan 2024
https://doi.org/10.5194/egusphere-egu23-15604
EGU General Assembly 2023

Forecasting discharges through explainable machine learning approaches at an alpine karst spring

Anna Pölz[1,4], Julia Derx[1,4], Andreas Farnleitner[2,3,4], and Alfred Paul Blaschke[1,4]
[1]Institute of Hydraulic Engineering and Water Resources Management, TU Wien, Vienna, Austria
[2]Institute of Chemical, Environmental and Bioscience Engineering, Research Group Microbiology and Molecular Diagnostics, TU Wien, Vienna, Austria
[3]Division Water Quality and Health, Karl Landsteiner University for Health Sciences, Krems, Austria
[4]Interuniversity Cooperation Centre Water and Health, Austria (www.waterandhealth.at)

Karst springs provide drinking water for approximately 700 million people worldwide. Complex subsurface flow processes lead to challenges for modelling spring discharges. Machine learning (ML) models possess the ability to learn non-linear patterns and show promising results in forecasting dynamic spring discharge. We compare the performance of three ML models of varying complexity in forecasting karst spring discharges: the multivariate adaptive regression spline model (MARS), a feed-forward neural network (ANN) and a long short-term memory model (LSTM). The well-studied alpine karst spring LKAS2 in Austria is used as test case. We provide model explanations including feature attribution through Shapley additive explanations (SHAP), a method based on Shapley values. Our results show that the higher the model complexity, the higher the accuracy, based on the evaluated symmetric mean absolute percentage error of the three investigated models. With SHAP every prediction can be explained through each feature in each input time step. We found seasonal model differences. For example, snow influenced the model mostly in winter and spring. Analyzing the combinations of input time steps and features provided further insights into the model performance. For instance, the SHAP results showed that a high electrical conductivity in recent time steps, which indicates that the karst water is less diluted with precipitation, leads to a reduced discharge forecast. These feature attribution results coincide with physical processes within karst systems. Therefore, the introduced SHAP method can increase the confidence in ML model forecasts and emphasizes the raison d’être of complex and accurate deep learning models in hydrology. This allows the operator to better understand and evaluate the model’s output, which is essential for drinking water management.