EGU23-11347, updated on 26 Apr 2024
https://doi.org/10.5194/egusphere-egu23-11347
EGU General Assembly 2023

Data-driven streamflow forecasting analysis leveraging multiple meteorological providers

Daniele Dalla Torre, Andrea Menapace, Ariele Zanfei, and Maurizio Righetti
Free University of Bolzano, Bolzano (BZ), Italy

In the last years, different providers (e.g. NOAA, MeteoFrance, ECMWF, DWD) produced meteorological gridded global and regional data sets. These model outputs are now reanalysis and operative forecasts distributed as open data, with different spatial and temporal resolutions. The aim of this contribution is the comparison of short-term streamflow forecasting outputs using different meteorological data sets as forcing.

The use of the above-predicted weather time series as input for the hydrological models allows the evaluation of the streamflow at the catchment closure. Hourly data sets of temperature and precipitation from the different providers were selected for this contribution and the evaluation of the short-term streamflow forecasting results on three small catchments in the Alps was carried out.

A data-driven forecasting procedure with the Support Vector Regression machine learning algorithm as a tool for hydrological modeling was implemented. For each provider, training and testing phase were performed using as forcing the weather model outputs of the same provider, in this way the simulations are consistent. The training phase takes advantage of reanalysis data sets when available, otherwise historical forecasting products are used as an alternative. Instead, the testing period is forced by lags of the temperature, precipitation and streamflow metered data for the past. The number of lag hours is defined in the training stage with a grid search approach. Moreover, the actual temperature and precipitation forecasting data sets cover the prediction lead time. The training was carried on for each day of the training period and the output of each run is the hourly short-term streamflow prediction.

The use of multiple data sources as input allows us to emphasize the differences between global and local meteorological forcing. Moreover, the simulation ensembles outputs allow the identification and quantification of uncertainty that lead to a better interpretation of the prediction. These results are promising in hydrological modeling to increase the final accuracy of the streamflow predictions and the decision-making under uncertainty.