EGU21-4567, updated on 17 Oct 2021
https://doi.org/10.5194/egusphere-egu21-4567
EGU General Assembly 2021

Testing different machine learning techniques for runoff routing in a highly glacierized Djankuat river basin (the North Caucasus, Russia).

Ekaterina Rets, Sergey Fomichev, and Egor Belozerov
Water Problems Institute Russian Academy of Sciences, Moscow, Russian Federation (info@iwp.ru)

Distributed, physically based modelling of runoff routing in highly glacierized river basins is an extremely complicated task as glacier drainage systems functioning is very sophisticated, close to karst river systems but also dynamically developing within very short time periods. Accordingly, runoff routing of glacier melt water is most often based on the concept of linear storage. The number of reservoirs generally vary from 1 to 3. For example, one 'fast' reservoir for melt and rain in glacierized grid cells in GERM model, three parallel different linear reservoirs representing snow, firn and ice in GSM-SOCONT model.

Here we test applicability of different machine learning techniques (gradient boosting, random forest, LSTM) for runoff routing in a highly glacierized river basin. We use the data from Djankuat alpine research catchment located in the North Caucasus (Russia) for the period of 2007-2019. The dataset contains different parameters measured with an hourly or sub-daily time step: water runoff, conductivity, turbidity, temperature, 18O, D content at the main gauging station; measurements of precipitation amount, standard meteorological parameters and radiation fluxes. Results of snow and ice melting modelling in the Djankuat river basin over a regular net with an hourly time step using energy-balance distributed A-Melt model are also used as input data.

Total runoff from the Djankuat river basin (1) and meltwater runoff according to isotopic hydrograph separation (2) were chosen as target functions. Different sets of features to predict the target functions were generated from the original time series using different combinations of the input parameters as well as variable lag times. To score different machine learning techniques and sets of features to predict target function we use correlation coefficient, Nash-Sutcliff efficiency index (NSE), root mean square error (RMSE).

The study was supported by the Russian Foundation for Basic Research, grant No. 20-35-70024