EGU2020-11553
https://doi.org/10.5194/egusphere-egu2020-11553
EGU General Assembly 2020

Objective classification of changes in water regime types of the Russian Plain rivers utilizing machine learning approaches

Alexander Ivanov[1], Timophey Samsonov[2], Natalia Frolova[1], Maria Kireeva[1], and Elena Povalishnikova[1]
[1]Lomonosov Moscow State University, Faculty of Geography, Land Hydrology department(al.m.ivanov@gmail.com)
[2]Lomonosov Moscow State University, Faculty of Geography, Cartography and Geoinformatics department (tsamsonov@geogr.msu.ru)

Hydrological regime classification of Russian Plain rivers was always done by hand and by using subjective analysis of various characteristics of a seasonal runoff. Last update to this classification was made in the early 1990s.

In this work we make an attempt at using different machine learning methods for objective classification. Both clustering (DBSCAN, K-Means) and classification (XGBoost) methods were used to establish 1) if an established runoff types can be inferred from the data using supervised approach 2) similar clusters can be inferred from data (unsupervised approach). Monthly runoff data for 237 rivers of Russian Plain since 1945 and until 2016 were used as a dataset.

In a first attempt dataset was divided into periods of 1945-1977 and 1978-2016 in attempt to detect changes in river water regimes due to climate change. Monthly data were transformed into following features: annual and seasonal runoff, runoff levels for different seasons, minimum and maximum values of monthly runoff, ratios of the minimum and maximum runoff compared to yearly average and others. Supervised classification using XGBoost method resulted in 90% accuracy in water regime type identification for 1945-1977 period. Shifts in water regime types for southern rivers of Russian Plain rivers in a Don region were identified by this classifier.

DBSCAN algorithm for clustering was able to identify 6 major clusters corresponding to existing water regime types: Kola peninsula, North-East part of Russian Plain and polar Urals, Central Russia, Southern Russia, arid South-East, foothills and separately higher altitudes of the Caucasus. Nonetheless a better approach was sought due to intersections of a clusters because of the continuous nature of data. Cosine similarity metric was used as an alternative way to separate river runoff types, this time for each year. Yearly cutoff also allows us to make a timeline of water regime changes over the course of 70 years. By using it as an objective ground truth we plan to remake classification and clusterization made earlier and establish an automated way to classify changes in water regime over time.

As a result, the following conclusions can be made

The study was supported by the Russian Science Foundation (grant No.19-77-10032) in *methods and Russian Foundation for Basic Research (grant No.18-05-60021) for analyses in Arctic region*