

EGU23-14210, updated on 21 Apr 2024  
<https://doi.org/10.5194/egusphere-egu23-14210>  
EGU General Assembly 2023  
© Author(s) 2024. This work is distributed under the Creative Commons Attribution 4.0 License.



## Detection of Arctic rivers streamflow drivers through automatic feature selection

**Mattia Zeno**, Matteo Sangiorgio, and Andrea Castelletti  
Politecnico di Milano, Milan, Italy ([mattia1.zeno@mail.polimi.it](mailto:mattia1.zeno@mail.polimi.it))

Runoff from Arctic rivers has a direct influence on the sea ice dynamics in the Arctic Ocean, producing significant effects from local to global scales. Despite their key role, the knowledge of the processes that influence the Arctic rivers streamflow is still limited, and their behavior is not fully understood.

In the literature, these analyses are usually performed adopting classical statistical methods and simple linear models, which are probably unable to fully capture underlying nonlinearities and redundancy of candidate drivers.

In this study, we use automatic feature selection techniques to detect the main drivers of the five major Arctic Rivers' runoff (Ob, Yenisei, and Lena in Asia, Mackenzie and Yukon in North America). Daily time series of temperature and precipitation recorded by several stations spread across the Arctic region, and the average snow cover of each basin are used as candidate input variables.

The feature selection analysis is carried out with two algorithms: Wrapper for Quasi Equally Informative Subset Selection (W-QEISS) and Iterative Input Selection (IIS). W-QEISS adopts neural predictive models to select alternative sets of drivers providing similar in terms of accuracy, but with different relevance, redundancy, and cardinality. Conversely, IIS directly produces a ranking of the input variables relying on tree-based models and combining computational efficiency and scalability to high input dimensionality.

The two algorithms achieve noticeably consistent results, with minor differences that can be explained by numerical factors typical of machine learning. Results also show that autoregressive terms have a crucial role in all the hydrological basins, while the importance of the other drivers is different for each river.

This preliminary research opens the floor for further analysis to broaden the knowledge of Arctic hydro-meteorological dynamics.