EGU24-14957, updated on 23 Jan 2025
https://doi.org/10.5194/egusphere-egu24-14957
EGU General Assembly 2024

Exploring data-driven emulators for snow on sea ice

Ayush Prasad[1,2], Ioanna Merkouriadi[3], and **Aleksi Nummelin**[4,5]
[1]University of Helsinki, Helsinki, Finland (ayush.prasad@helsinki.fi)
[2]Finnish Meteorological Institute, Helsinki, Finland (ayush.prasad@fmi.fi)
[3]Finnish Meteorological Institute, Helsinki, Finland (ioanna.merkouriadi@fmi.fi)
[4]Finnish Meteorological Institute, Helsinki, Finland (aleksi.nummelin@fmi.fi)
[5]NORCE Norwegian Research Centre AS, Bergen, Norway

Snow is a crucial element of the sea ice system, impacting various environmental and climatic processes. SnowModel is a numerical model that is developed to simulate the evolution of snow depth and density, blowing-snow redistribution and sublimation, snow grain size, and thermal conductivity, in a spatially distributed, multi-layer snowpack framework. However, SnowModel faces challenges with slow processing speeds and the need for high computational resources. To address these common issues in high-resolution numerical modeling, data-driven emulators are often used. They aim to replicate the output of complex numerical models like SnowModel but with greater efficiency. However, these emulators often face their own set of problems, primarily a lack of generalizability and inconsistency with physical laws. A significant issue related to this is the phenomenon of concept drift, which may arise when an emulator is used in a region or under conditions that differ from its training environment. For instance, an emulator trained on data from one Arctic region might not yield accurate results if applied in another region with distinct snow properties or climatic conditions. In our study, we address these challenges with a physics-guided approach in developing our emulator. By integrating physical laws that govern changes in snow density due to compaction, we aim to create an emulator that is efficient while also adhering to essential physical principles. We evaluated this approach by comparing four machine learning models: Long Short-Term Memory (LSTM), Physics-Guided LSTM, Gradient Boosting Machines, and Random Forest, across five distinct Arctic regions. Our evaluations indicate that all models achieved high accuracy, with the Physics-Guided LSTM model demonstrating the most promising results in terms of accuracy and generalizability. This approach offers a computationally faster way to emulate the SnowModel with high fidelity.