

Accelerating Subglacial Bed Topography Prediction in Greenland: A Performance Evaluation of Spark-Optimized Machine Learning Models

Accurate estimation of subglacial bed topography is crucial for understanding ice sheet dynamics and their responses to climate change. In this study, we employ machine learning models, enhanced with Spark parallelization, to predict subglacial bed elevation using surface attributes such as ice thickness, flow velocity, and surface elevation. Radar track data serves as ground truth for model validation. Our primary objective is to leverage Spark's distributed computing framework to accelerate model training and evaluation, enabling scalable analysis of large datasets. We tested several machine learning algorithms compatible with Spark, including XGBoost, Gradient Boosting (GBoost), Random Forest, and Kernel Regression. XGBoost emerged as the most efficient model, achieving substantial speed-ups as the number of computing nodes increased. Our findings underscore the importance of distributed computing in enhancing the scalability and efficiency of machine learning models for large-scale climate data. The transition to the High-Performance Computing Facilities (HPCF), along with Spark parallelization, significantly reduced training time, demonstrating the effectiveness of distributed computing for complex datasets. This approach not only improves computational performance but also accelerates experimentation and analysis, contributing to a deeper understanding of ice sheet behavior and its implications for climate change. Future work will focus on applying these methods to even larger datasets, further leveraging Spark's capabilities to advance predictive modeling and support climate change mitigation and adaptation efforts. Moreover, we plan to explore deep neural networks and their performance on this application, leveraging multi-GPU architectures to further accelerate model training.

Focus areas

Authors: CHAM, Mostafa (iHARP); TABASSUM, Tartela (UMBC); SHAKERI, Ehsan (UMBC); Dr WANG, Jianwu (iHARP)