EGU2020-22587
https://doi.org/10.5194/egusphere-egu2020-22587
EGU General Assembly 2020

Deep neural networks for total organic carbon prediction and data-driven sampling

Everardo González Ávalos and Ewa Burwicz
GEOMAR Helmholtz Centre for Ocean Research, Kiel, Germany (egonzalez@geomar.de)

Over the past decade deep learning has been used to solve a wide array of regression and classification tasks. Compared to classical machine learning approaches (k-Nearest Neighbours, Random Forests,...) deep learning algorithms excel at learning complex, non-linear internal representations in part due to the highly over-parametrised nature of their underling models; thus, this advantage often comes at the cost of interpretability. In this work we used deep neural network to construct global total organic carbon (TOC) seafloor concentration map. Implementing Softmax distributions on implicitly continuous data (regression tasks) we were able to obtain probability distributions to asses prediction reliability. A variation of Dropout called Monte Carlo Dropout is also used during the inference step providing a tool to model prediction uncertainties. We used these techniques to create a model information map which is a key element to develop new data-driven sampling strategies for data acquisition.