EGU24-9731, updated on 27 Nov 2024
https://doi.org/10.5194/egusphere-egu24-9731
EGU General Assembly 2024

Soil Organic Carbon Retrieval from DESIS Images by CNN

Xiangyu Zhao[1], Uta Heiden[2], Paul Karlshöfer[2], Zhitong Xiong[1], and Xiao Xiang Zhu[1]
[1]Technical University of Munich, School of Engineering and Design, Chair of Data Science in Earth Observation, Munich, Germany
[2]German Aerospace Center (DLR),The Remote Sensing Technology Institute, Wessling, Germany

Imaging spectroscopy is commonly used for many applications like soil, water, and vegetation. Digital soil mapping, especially by space-borne sensors, has become advantageous and promising due to its high efficiency. By using multispectral or hyperspectral images, topsoil properties could be estimated efficiently and accurately on a large area scale. Moreover, deep learning has been explored in the remote sensing community and achieved excellent performances in many remote sensing tasks. In this work, we explore deep learning methods to retrieve Soil Organic Carbon (SOC) value from DESIS hyperspectral images for the whole Bavaria state in Germany. For the hyperspectral data, we use all available DESIS images in Bavaria, which is 560 in total. Regarding the soil data, we combine SOC data from LFU (Bavarian State Environment Agency) and LUCAS 2018 (Land Use and Coverage Area frame Survey). Following a rigorous data selection process, we opted to include 1200 soil samples in our experiments. Starting from the raw hyperspectral images, we conduct a few preprocessing steps such as land cover masking, filtering by NDVI, building temporal composite, and then extracting patches surrounding each soil sample. These preprocessed patches are fed into deep learning models such as 1D CNN and 2D CNN, which are trained to predict the SOC value. To better interpret the model's performance, we also compute the SHAP(Shapley Additive Explanations) value for both frameworks. Specifically, we explore the SHAP value in spectral dimension for 1D CNN and analyze digital elevation features with 2D CNN in spatial dimension. During experiments, we split the whole dataset into train, validation, and test. To evaluate the performance, RMSE, R^2, and RPID are computed. For the specific structure of the models, many different parameters are investigated in parameter tuning. For each trial, 5 cross-validation is applied. In the end, we visualize the prediction results by a soil map. From the results, the best-performed model could get RMSE 0.62 and R^2 0.40 on the test set. Moreover, we find that the first-order derivative of the spectrum is the most important feature for predicting SOC, while 1D CNN is capable of extracting useful information from it and achieving excellent regression results with RMSE 0.66 and R^2 0.32. Additionally, spectrums between 530 nm - 570 nm and 730 nm - 780 nm are the most informative according to SHAP analysis.