Predicting topsoil organic carbon using UAV-based hyperspectral sensor

János Mészáros¹, Gergely Jakab², Mátyás Árvai¹, Judit Szabó¹, Márton Tóth¹, Boglárka Keller³, Gábor Szatmári¹, Zoltán Szalai², and László Pásztor¹
¹Department of Soil Mapping and Environmental Informatics, Institute for Soil Sciences and Agricultural Chemistry, Budapest, Hungary (meszaros.janos@rissac.hu)
²Geographical Institute, Research Centre for Astronomy and Earth Sciences, Budaörsi út 45, 1112 Budapest, Hungary
³Dalmand Zrt, Bonafarm Group, Felszabadulás utca 42, 7211 Dalmand, Hungary

There is increasing demand for up-to-date spatial information on soil organic carbon (SOC). Meanwhile, Unmanned Aerial Vehicles (UAV) provide flexible technology for monitoring land surface features with high spatial resolution at plot scale. Suitably performed, airborne imagery simultaneously provides spectral and terrain based spatial auxiliary data, which can be used as predictors in DSM-type modelling of topsoil OC.

To test its applicability for spatial prediction of topsoil OC, an aerial survey was carried out on a plot situated on a gently undulating slope by a Cubert UHD-185 hyperspectral snapshot camera mounted on a Pixhawk-based octocopter. The camera is capable to record electromagnetic spectrum between 450-950 nm in 125 spectral bands on 50×50 pixels images and the panchromatic spectrum in 1 Mpx images. Because of the narrow field-of-view of the UHD-185, three consecutive flights were needed to cover the whole area (cca. 10 ha); all were happened in the hours close to noon and flown in automatic flight mode to ensure the right over- and sidelap between images to make possible the photogrammetric processing. Despite the automatic flights a surveying grade GPS unit was also used to survey 12 markers, evenly distributed on the field to orthorectify images later.

The hyperspectral and panchromatic images were pre-processed in Cubert Edelweiss to produce different versions of them depending on the used spectral information to investigate later how built-in pan-sharpening method affects the prediction accuracy. The generated datasets are the native and pan-sharpened hyperspectral mosaics. Later the photogrammetric processing was performed in Agisoft Photoscan for both hyperspectral datasets, resulting in two georeferenced outcomes: a common digital elevation model (DEM) and two hyperspectral orthomosaics of the area, each exported with 1 m spatial resolution. Further data editing steps were carried out in R, generating various versions of exported hyperspectral orthomosaics: mosaic containing all of the 125 spectral bands; filtered (where spectrally overlapping bands with high correlation were removed based on Full Width at Half Minimum information) and Principal Component Analysis transformed versions.
Based on different kinds of spectral orthomosaics and DEM combinations, a custom R script using Random Forest algorithm generated 36 predicted layers for topsoil OC, which were validated by Leave-One-Out Cross-Validation, hence independent mean and RMSE errors could be calculated for each dataset combinations. The overall best performing datasets were provided by the FWHM-filtered hyperspectral orthomosaic, hence the lowest mean error is resulted by the filtered, pan-sharpened PCA-transformed combination containing the DEM and its derivatives. However, in the RMSE values there were no significant difference between the six lowest RMSE combinations, but mostly the pan-sharpened and PCA-transformed versions perform better.