



## **Within-field and regional-scale accuracies of topsoil organic carbon content prediction from an airborne visible near-infrared hyperspectral image combined with synchronous field spectra for temperate croplands**

Emmanuelle Vaudour (1), Jean-Marc Gilliot (1), Liliane Bel (2), Josias Lefevre (3), and Kacem Chehdi (3)

(1) UMR ECOSYS, AgroParisTech, INRA, Université Paris-Saclay, 78850, Thiverval-Grignon, France (emmanuelle.vaudour@agroparistech.fr), (2) UMR MIA-Paris, AgroParisTech, INRA, Université Paris-Saclay, 75005, Paris, France, (3) ENSSAT/TSI2M, 6 rue de Kérampont, B.P. 80518, F-22305 Lannion cedex, France

This study was carried out in the framework of the TOSCA-PLÉIADES-CO of the French Space Agency and benefited data from the earlier PROSTOCK-Gessol3 project supported by the French Environment and Energy Management Agency (ADEME). It aimed at identifying the potential of airborne hyperspectral visible near-infrared AISA-Eagle data for predicting the topsoil organic carbon (SOC) content of bare cultivated soils over a large peri-urban area (221 km<sup>2</sup>) with intensive annual crop cultivation and both contrasted soils and SOC contents, located in the western region of Paris, France. Soils comprise hortic or glossic luvisols, calcaric, rendzic cambisols and colluvic cambisols.

Airborne AISA-Eagle images (400-1000 nm, 126 bands) with 1 m-resolution were acquired on 17 April 2013 over 13 tracks. Tracks were atmospherically corrected then mosaicked at a 2 m-resolution using a set of 24 synchronous field spectra of bare soils, black and white targets and impervious surfaces. The land use identification system layer (RPG) of 2012 was used to mask non-agricultural areas, then calculation and thresholding of NDVI from an atmospherically corrected SPOT4 image acquired the same day enabled to map agricultural fields with bare soil. A total of 101 sites, which were sampled either at the regional scale or within one field, were identified as bare by means of this map.

Predictions were made from the mosaic AISA spectra which were related to SOC contents by means of partial least squares regression (PLSR). Regression robustness was evaluated through a series of 1000 bootstrap data sets of calibration-validation samples, considering those 75 sites outside cloud shadows only, and different sampling strategies for selecting calibration samples. Validation root-mean-square errors (RMSE) were comprised between 3.73 and 4.49 g. Kg<sup>-1</sup> and were  $\sim 4$  g. Kg<sup>-1</sup> in median. The most performing models in terms of coefficient of determination ( $R^2$ ) and Residual Prediction Deviation (RPD) values were the calibration models derived either from Kennard-Stone or conditioned Latin Hypercube sampling on smoothed spectra. However, the most generalizable model leading to lowest RMSE value of 3.73 g. Kg<sup>-1</sup> at the regional scale and 1.44 g. Kg<sup>-1</sup> at the within-field scale and low validation bias was the cross-validated leave-one-out PLSR model constructed with the 28 near-synchronous samples and raw spectra.