# MAPPING SMALLHOLDER AGRICULTURE USING SIMULATED SENTINEL-2 DATA; OPTIMIZATION OF A RANDOM FOREST-BASED APPROACH AND EVALUATION ON MADAGASCAR SITE

Lebourgeois V[1], Dupuy S[1], Vintrou E[1], Ameline M[1], Butler S[1], Begue A[1]
[1]Cirad

Thanks to the recent launch of Sentinel-2 satellite, access to a fine monitoring of the crops is now possible for smallholder agricultural systems. However, many constraints still need to be addressed, such as time-series incompleteness in cloudy environments, high intra- and inter-fields spatial variability, and large volume of data.

In this paper, we propose 1. to test a combined Random Forest (RF) classifier applied to multisource satellite data with an object-based approach, in order to produce land use maps at different levels (from cropland to crop type) ; 2. to optimize the performance of the RF classifier by reducing the number of input variables. The test site is a small agricultural zone in Madagascar. RF is applied by using two approaches: classical (by classifying the different levels from the whole learning database) and hierarchical (by classifying first the cropland and non-cropland and classifying crop types inside the cropland only). The importance (mean decrease accuracy measure) of the different variables is analyzed according to their source (high spatial resolution (HRS) time series simulating Sentinel-2 data, PLEIADES images, DEM), and category (reflectance, texture, spectral indices, auxiliary data). The contribution to the classification accuracy of each source andcategory of variables is also tested.

Results showed that optimizing the RF classifier allowed dividing the number of variables to be extracted by 1.5 to 6, depending on the classification level. Using a hierarchical approach improved the classification results whatever the level of classification and allowed to obtain 64.4% overall accuracy for crop type classification. The HRS time series derived spectral variables revealed to be the more discriminant, with a slight advantage of spectral indices over reflectances. Next step is the application of the optimized approach to a Sentinel-2 time series acquired over a large area in the context of the Sen2-Agri demonstration phase.