ON THE EFFECT OF SAMPLING SCHEMES ON CROPLAND CLASSIFICATION ACCURACY

F. Waldner et al.
October 11th 2016, Kiev
A FIRST SUCCESSFUL EXPERIMENT

The site-effect is stronger than the method-effect
→ Need for site-specific methods
A SECOND EXPERIMENT TO SUPPORT THE GUIDELINES FOR IN SITU DATA COLLECTION

“The overarching goal of JECAM is to reach a convergence of approaches, develop monitoring and reporting protocols and best practices for a variety of global agricultural systems. JECAM will enable the global agricultural monitoring community to compare results based on disparate sources of data, using various methods, over a variety of global cropping systems.”

From the JECAM Guidelines for cropland and crop type definition and field data collection
THE EFFECT OF THE SAMPLING DESIGN

The overarching objective is to assess the ability of sampling schemes to deliver:

1. Accurate **cropland maps**;
2. Reliable **accuracy indicators**.

Assessment over multiple sites with large field size (open to anyone interested)

Natural vegetation gradients; spatial variations in management practices; soils; other land cover classes.

→ Argentina, Brazil, China, Ukraine, Russia

→ 200x200km²
TWO MAIN OBJECTIVES

Two effects to jointly investigate:

1. the **sampling design**
2. the **number of samples**

**CALIBRATION**

Accuracy assessment of the cropland maps

Assessment of thematic certainty resulting from the classification

**VALIDATION**

Introduce misclassification on purpose and assess the ability of a sampling scheme to give a reliable accuracy estimate
CASE STUDY IN RUSSIA BASED ON A PRELIMINARY PROTOCOL

October 11th 2016, Kiev
METHODOLOGY

Satellite data

Parcel data

Sampling scheme

Calibration data

SVM classifier

Thematic certainty

Croppland map

Satellite data: annual MODIS time-series
SETTING UP THE SANDBOX

1. Download the available and cloud-free Sentinel-2 data (red, nir)

2. Segmentation of the imagery and interactive supervised classification to extract the fields.

3. Manual correction by the field partners

3. Resampling to the spatial resolution of interest (in this case MODIS).
DIFFERENT SAMPLING STRATEGIES

**Systematic;**

**Random;**

**Area frame-like;**

**Random stratified;**

Using GL30 classes as strata

**Along roads;**

Buffer areas Based on level-2 OpenStreetMap

...
SAMPLES USED AS CALIBRATION DATA — ACCURACY

with 600 calibration points
SAMPLES USED AS CALIBRATION DATA — THEMATIC CONFIDENCE
SAMPLES USED AS VALIDATION DATA

1. Random generation of maps with known accuracy;

2. Assessment of the reliability of the accuracy estimate given by each sampling scheme