RAF/AFCAS/29/8.1 E

November 2025

منظمة الأغذية والزراعة للأمم المتحدة | 联合国粮食及农业组织 | Food and Agriculture Organization of the United Nations | Organisation des Nations Unies pour l'alimentation et l'agriculture | Продовольственная и сельскохозяйственная организация Объединенных Наций | Organización de las Naciones Unidas para la Agricultura y la Alimentación

AFRICAN COMMISSION ON AGRICULTURAL STATISTICS

Twenty-Nineth Session

(Hammamet, Tunisia)
24 – 28 November 2025

Gap Analysis and Solutions on EO Requirements and Field Data Collection Protocols in Agricultural Surveys: Case Studies from Senegal and Uganda

Executive summary

This paper presents a gap analysis and solutions to improve alignment between field data collection protocols in national agricultural surveys and the requirements of Earth Observation (EO) applications. Drawing from case studies in Senegal and Uganda under the FAO-led EOSTAT and ESA-Sen4Stat initiatives, the study introduces a practical framework to assess the fitness-for-use of in-situ data for EO-based agricultural statistics. In Senegal, aligning protocols with EO needs led to a classification accuracy increase from 78% to over 85% and allowed combining the agricultural survey with this map to decrease the variance associated with the acreage estimates. In Uganda, results showed that solid spatial design and georeferencing, but gaps in semantic accuracy and post-hoc EO validation (e.g., no NDVI/spectral purity checks) can be workable (≈78% OA) but produce less reliable classification and area indicators that required heavy error corrections.

The findings underscore the importance of redesigning agricultural surveys to meet EO standards and of phased testing to ensure institutional uptake.

Suggested actions by AFCAS

The Commission is invited to:

1. Promote the adoption of in-situ data quality framework within national agricultural survey programmes and consider FAO's in-situ quality framework in this regard.
2. Assess their existing survey protocols using diagnostic criteria proposed by FAO (table 1)
3. Support capacity development on EO-based quality assessment techniques.
4. Consider national pilots to test improved protocols prior to nationwide adoption.

Queries on the substantive content of this document may be addressed to:
AFCAS Secretariat;

Email: dominique.habimana@fao.org

1. Introduction

Earth Observation (EO) technologies are increasingly recognized as transformative tools for enhancing agricultural statistics. However, their integration into official systems is often constrained by the structure and quality of in-situ data collected via agricultural surveys and censuses.

In response, FAO, through the EOSTAT programme and European Space Agency Sen4Stat projects, developed a diagnostic framework to assess the fitness-for-use of collected in-situ data for EO applications. This paper introduces this framework and applies it to two country case studies—Senegal and Uganda—to illustrate common challenges and propose actionable solutions.

2. FAO in-situ quality framework

FAO and Université catholique of Louvain in Belgium (UCLouvain) have been testing and experimenting data quality framework for the collection of in situ data that would ensure fitness for use with EO applications, in particular to combine the national surveys with crop type maps with the finality to improve crop statistics.

Finally, in 2025, the experience gained after several field experiments culminated in the publishing of a first in-situ quality framework for assessing the compatibility of field data collected through Surveys and Census operations with EO applications (De Simone, L. et al. 2025[1]).

The framework includes two pillars:

- Survey Design Assessment (**Table 1**): Evaluates the structural compatibility of sampling protocols with EO needs (e.g., crop observability, georeferencing precision, and metadata completeness).
- EO-Based Post-Hoc Assessment (**Table 2**): Applies satellite-derived analytics (e.g., Normalized Difference Vegetation Index time series clustering, visual cross-checks) to validate the internal consistency and label quality of field samples.

While the framework does not assign numerical scores, it identifies key design gaps and recommends corrective actions. A dataset is considered unfit if significant gaps in either pillar remain unresolved.

[1] De Simone L, Pizarro E, Paredes J, et al. Quality control of training samples for agricultural statistics using Earth observation. *Statistical Journal of the IAOS*. 2025;41(2):467-481. doi:10.1177/18747655251338033

Table 1. Assessment framework to qualify the compatibility of an in-situ survey design to leverage EO satellite data for agriculture statistics
***1 Criteria related to the sampling design*:**
1.1 Observation timing allowing identification of crop type in the field (the survey must take place when the crop is visible in the field)
1.2 Minimum number of samples for marginal crops (including intercrop types) to provide balanced datasets in terms of crop type sample distribution
1.3 Local homogeneity of each sample unit to match the satellite observation footprint
2 Criteria related to the response design:
2.1 Georeferenced ground observation at field or point level to link with satellite geospatial dataset (household geographic coordinates being insufficient)
2.2 Sample unit size at least matching the satellite observation footprint
2.3 Contextual observation to document sample quality and qualify its representativity
2.4 Rich labelling of each sample beyond crop type to indicate specific growing conditions (e.g., weeds abundance, limited crop cover, water lodging, tree density, etc.)

Source: De Simone, L. et al. 2025 Doi:10.1177/18747655251338033

Table 2. EO-based assessment framework to control the quality of an in-situ dataset from a satellite remote sensing point of view
3 Criteria for EO-based data quality control:
3.1 Completeness of sample information , including rich labels linked with geospatial features and survey data
3.2 Topology integrity of georeferenced features collected on the ground (e.g. field polygon, roadside)
3.3 Metadata completeness including geographic coordinates system definition, protocol of geodata collection, and typology definition of all crop labels and related attributes
3.4 GPS coordinate accuracy visually checked with EO imagery
3.5 Sample purity based on multispectral reflectance heterogeneity assessment3.6 Labelling quality based on NDVI time series clustering and profile similarity analysis

Source: De Simone, L. et al. 2025 Doi:10.1177/18747655251338033

3. Case Study 1: Senegal

Since 2019, FAO and UCLouvain have collaborated with the Direction de l'Analyse, de la Prévision et des Statistiques Agricoles (DAPSA) under the EOSTAT project to assess and enhance the compatibility of Senegal's agricultural survey system with EO applications. The Annual Agricultural Survey (AAS) in place in Senegal is a list frame survey covering the whole country. Holdings are selected through a stratified sampling from the 526.000 holdings in Senegal, which corresponds approximately to an overall 0.4% sampling rate. These holdings were selected across with selection probability proportional to the size of these departments.
The AAS dataset from 2018 was shared by DAPSA as a baseline to evaluate its fitness for EO-based crop mapping using FAO's in-situ quality framework.

3.1 AAS 2018

The AAS dataset included 16,861 records corresponding to individual parcels cultivated by approximately 4,693 households surveyed across the country. In each holding, farmers were interviewed during the main cropping season and GPS measurements were done in a selection of fields, which were described in terms of crop type and crop practices (Figure 1). Parcels' areas were also measured with GPS and each measured areas was then reported manually in the survey questionnaire available on the tablet. Unfortunately, the GPS traces were not systematically archived – only a subset of them were kept for quality control by the supervision team at DAPSA. A subset of these traces, mainly located between Touba and Thiès, was also shared by DAPSA (Figure 2).

To summarize, for the AAS 2018, georeferenced information is available at crop-level, as points (parcels boundaries are not available by-default), for a selection of parcels. Information was collected at the right timing to allow crop identification and additional information

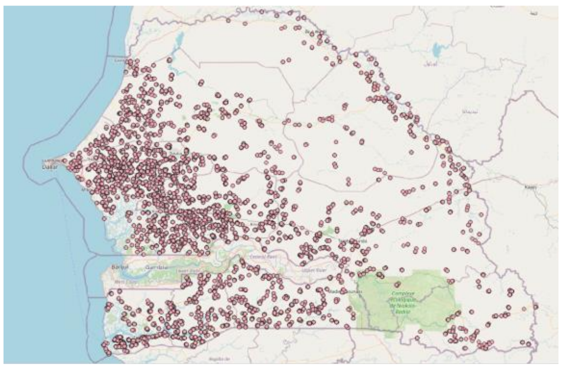


Figure 1: Household locations covered by the 2018 agricultural survey

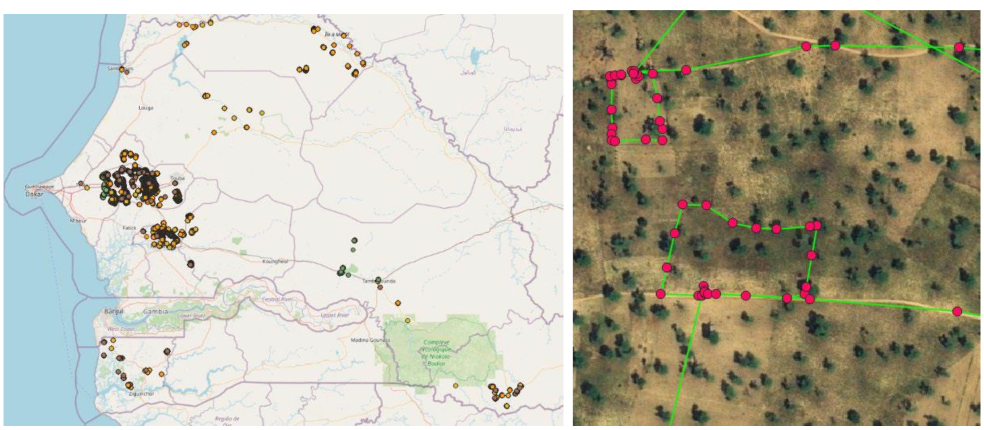


Figure 2. GPS tracks location (left) and parcels boundaries with GPS point/lines (right)

Quality control of AAS data

The quality of the 16,861 records (GPS points within the AAS database) was assessed through multiple stages. First, all points located too close to roads and buildings as well as points located less than 20 meters apart but associated with different crops were removed.

The second stage of the quality control addressed the important challenge of working with GPS points (and not polygons), i.e. checking if the point was located in the middle of the parcel. Figure 3 illustrates two situations where (i) point was taken at the middle of the parcels, thus making the parcels easily identifiable (left) and (ii) point was taken on the path along the parcels, making it almost impossible to know which parcel it refers to (right).

Figure 3. Challenge to localize the surveyed parcels in the statistical database due to the fact that parcels are localized by geo-points (i.e. a single GPS coordinate)

In order to avoid manually checking all points, the points were transformed into polygons by applying a 20-m buffer and the homogeneity of these polygons was assessed statistically. The assumption was that if the point was taken in the middle of the parcel, the resulting buffer would be homogeneous. Conversely, a point taken along the edge of the plot would result in a buffer mixing different crop types or including other land cover types such as roads or trees. This assessment was done calculating the standard deviation of the Normalized Difference Vegetation Index standard deviation.

During this step, it was observed that most points labelled as "rice" were located at the edges of the fields (probably because it is difficult to walk to the center of a rice field) and would thus be removed using the above criterion. In order to keep a good representation of rice in the reference database, all points were replaced manually in the middle of the fields thanks to Google Earth imagery. Following this stage, 12,056 polygons out of the 16,861 initial records remained in the sample.

Quality control of GPS traces

First, the raw GPS traces were converted into usable polygon shapefiles. This process involved removing irrelevant data, correcting naming inconsistencies, and cleaning line artifacts caused by enumerator movement between fields (Figure 4).

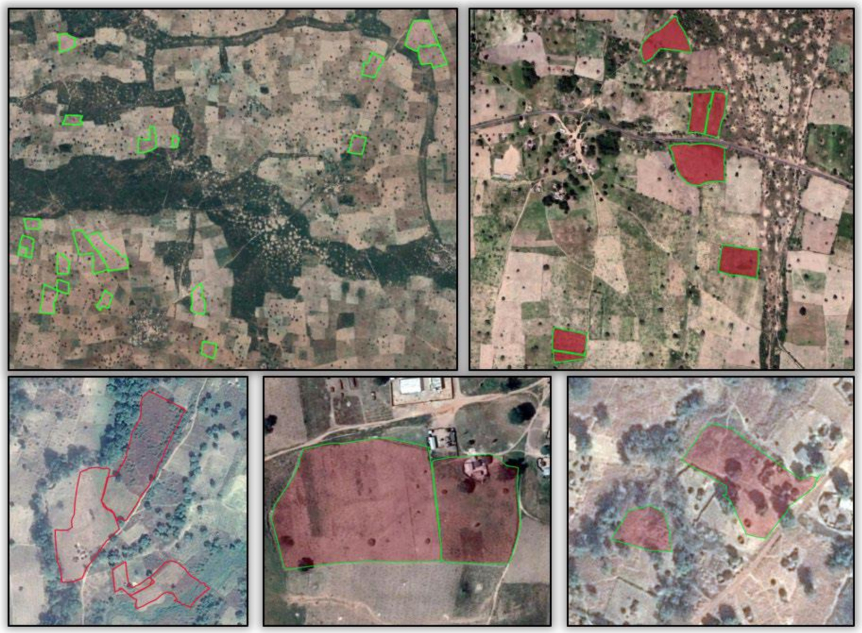


Figure 4. GPS traces successfully converted in shapefiles and represented well parcels boundaries

The GPS traces needed also to be quality controlled and filtered:

- Traces unrelated to valid records in the AAS database (e.g., invalid file name) were discarded;
- Traces including non-boundary movements (e.g., travel paths) were removed (Figure 5);
- Traces with fewer than 3 points were excluded as they could not form polygon.

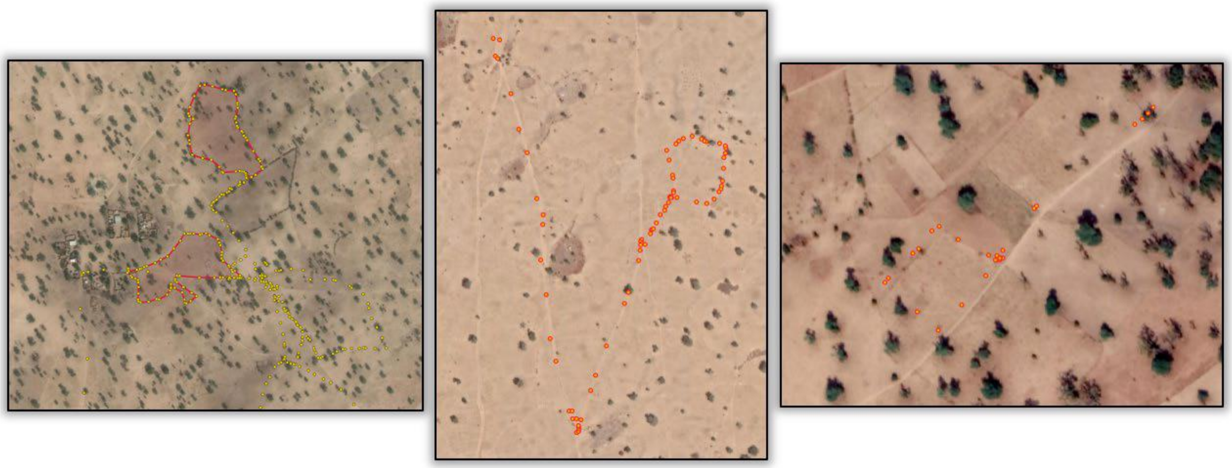


Figure 5. Example of GPS traces that contain more than parcels boundaries

Finally, too small polygons (the ones that disappeared after a 10-m inside buffer) were removed.

Ultimately, only 1,593 high-quality parcel boundaries remained, concentrated in central and western Senegal and representing 2089 hectares shared unequally into 22 crop type classes (Figure 6).

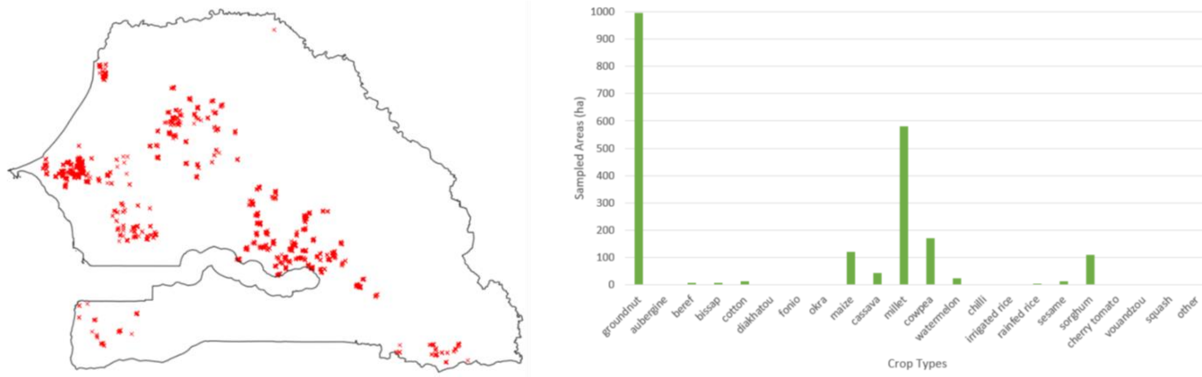


Figure 6. Localization of the 1593 crop polygons obtained from the GPS boundaries and distribution of crop polygons superficies according to the crop type

These crop data are expected to be the most useful ones because (i) they are polygons, thus including several pixels belonging to the same crop and (ii) the parcels boundaries have been checked manually. Nevertheless, these data don't contain enough samples in all the crop types and they are not spread over the whole country.

Crop mapping based on AAS 2018

Before running a classification using satellite imagery, additional non-crop information was collected by visual interpretation of very high spatial resolution Google Earth and Planet imagery. This information was not collected by regular agricultural surveys, which essence focuses on crop areas, but is needed to train classification algorithms.

The classification was run using both the points and polygons dataset, based on GPS traces. The class imbalance in the dataset was addressed by limiting the number of groundnut samples to a maximum of 400 hectares and of rice samples to 100 hectares.

These data enabled the generation of a national 10-meter spatial resolution binary cropland – non-cropland mask, with an overall accuracy of 98% and the F-Score values for the cropland and non-cropland classes of 76% and of 99%, respectively. A visual inspection of the map revealed that it performed relatively well: the discrimination was well done between cropland and the natural shrub and tree vegetation, the urban areas and the bare soil. The discrimination with the grassland was however of lower quality and the irrigated perimeters were not well identified as cropland, due to their poor representation in the in-situ data.

A national crop type map was also generated, focusing only on the main crops for which enough samples were available (groundnut, maize, cassava, millet, cowpea, rice and sorghum). The overall accuracy was 78%. The patterns of the fields were generally well identified, showing that the Sentinel-2 10-meter spatial resolution has the capacity of mapping crop types at field-level. However, the map was affected by significant confusion between crops.

A local crop type map, focusing only on the Sentinel-2 tile where polygons are available, was much more promising, with an overall accuracy increasing to 85% (Figure 7). It showed the added-value of working with polygons instead of points and of relying on many more samples to calibrate the algorithm.

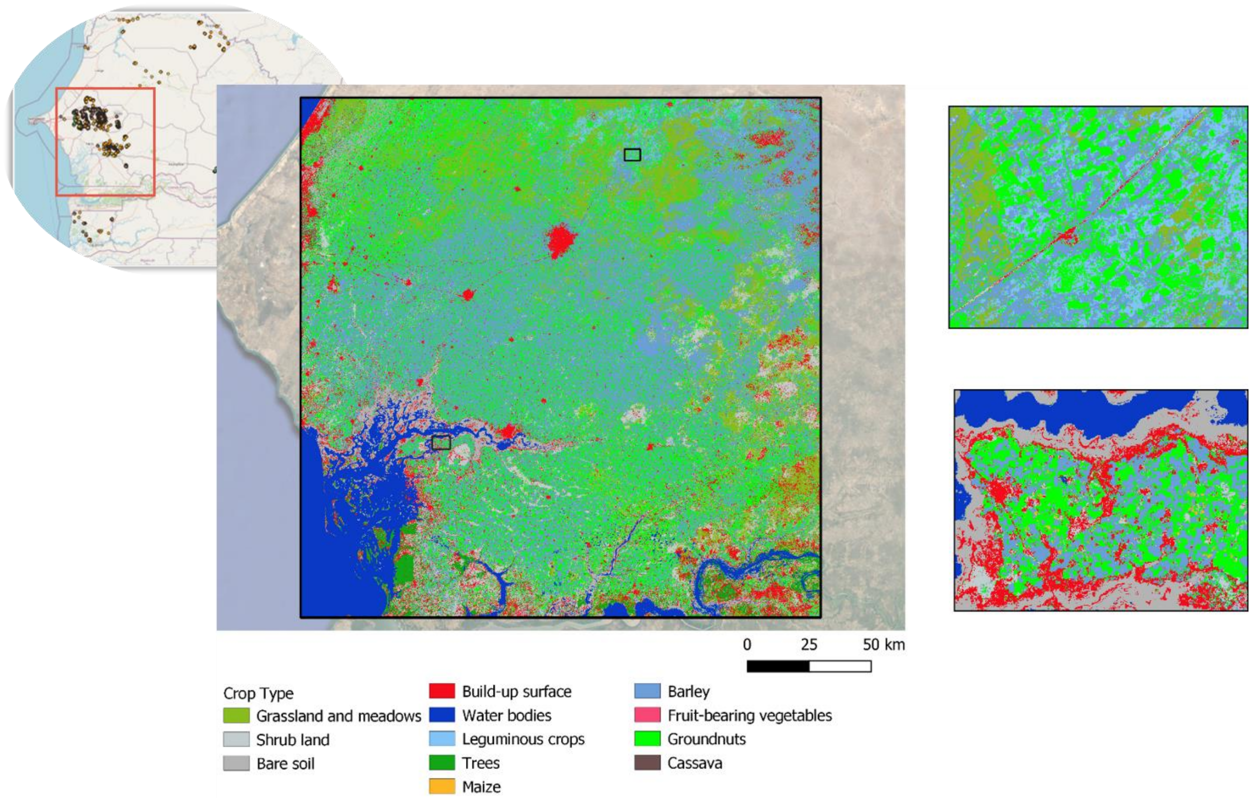


Figure 7. Sentinel-2 10-meter resolution crop type map over the tile where most of the GPS traces were located

The rigorous preprocessing and quality assessment of the AAS and GPS traces data played a pivotal role in achieving this result. However, the AAS protocol in place was not optimal to allow reaching higher mapping accuracy: GPS information was mainly recorded as points, minor crops were under-represented and contextual information was missing (e.g. field heterogeneity, mixed cropping).

3.2 Development and testing of an adjusted AAS protocol in 2021 in the Nioro department

Building on the lessons from the 2018 AAS, FAO supported a pilot activity in Nioro du Rip district between 2021–2022 to align the agricultural survey protocol with EO requirements. The proposed adjustments primarily targeted the georeferencing of the collected data by (1) transitioning from points to polygons and (2) testing area measurement using a tablet instead of a GPS, in order to simplify protocol implementation and reduce naming errors. It also added contextual and labelling information about crops in the questionnaire.

The pilot protocol discussed with the DAPSA consisted in recording parcels boundaries and no more GPS points for the visited parcels. These boundaries were recorded using two different devices for the sake of comparison: the traditional GPS and the tablet via a data collection application (which was ODK). The ODK application also allowed to collect additional information about crops, such as crop heterogeneity, weeds level and mixed cropping), as well as non cropland classes. The traditional AAS was implemented in parallel on the tablet as usual. Five teams worked on the field and the data collected were sent daily to a server accessible by all protagonists.

Simultaneously, DAPSA and UCLouvain carried out rapid quality control of the data each day in order to adjust the survey in real time if necessary.

Parcels boundaries with GPS vs ODK

GPS data
A total of 357 Garmin GPS traces were successfully recorded during the field campaign, compared to the 396 plots listed in the AAS database. The discrepancy was explained by two main issues:

- Plot naming inconsistencies: 29 polygons were not correctly linked to AAS data due to manual errors in GPS file naming. However, several polygons could be recovered through careful manual inspection.
- Loss of GPS traces: some GPS traces were never saved or transferred, preventing linkage with corresponding survey records.

These gaps in AAS–GPS linkage highlight the limitations of relying on manual data entry without real-time synchronization tools.

ODK data
A total of 231 plots were investigated in the ODK application, out of the 396 plots inventoried by the AAS. This difference was due to a communication error during data collection: taking measurements with two different tools on all parcels complicated the task for the investigators.

Taking the GPS traces as reference, the quality of the field boundaries collected with ODK was quite heterogeneous at the beginning of the field campaign (Figure 7). This was due to the time required to learn a new tool, and it was gradually resolved over the course of the campaign thanks to the continuous guidance provided by DAPSA and UCLouvain.

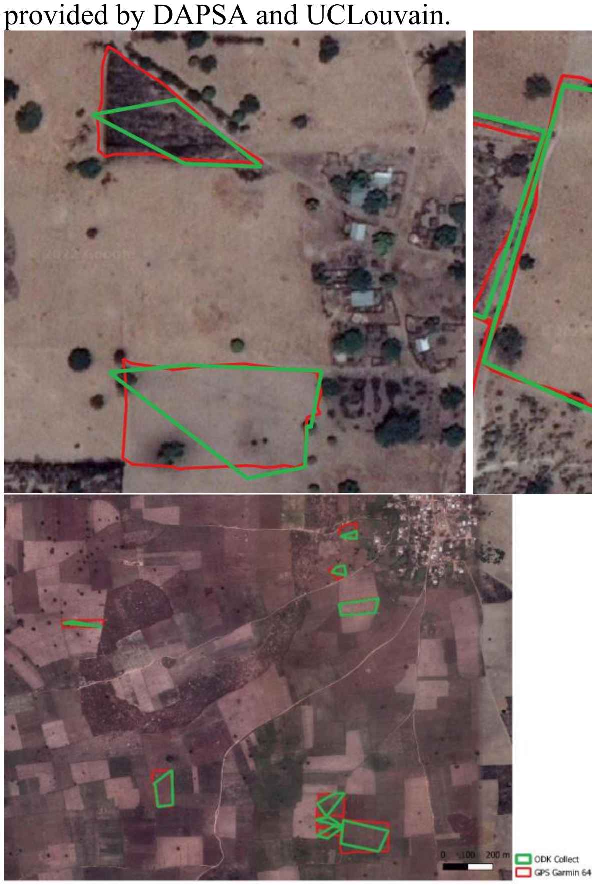


Figure 7: Parcel's polygons from Garmin GPS in red and from the tablet's GPS in green.

Comparative analysis of the surfaces
The data collected in the field by the tablet in manual mode were compared to the GPS data. Figure 8 relates the regression line of the tablet dataset to the reference line. The value of the slope is 0.99, close to the value of the reference slope. The position of the regression line and the bias of 0.27 indicates that the tablet data generally underestimated the reference data. This bias is confirmed by a visual analysis of the data.

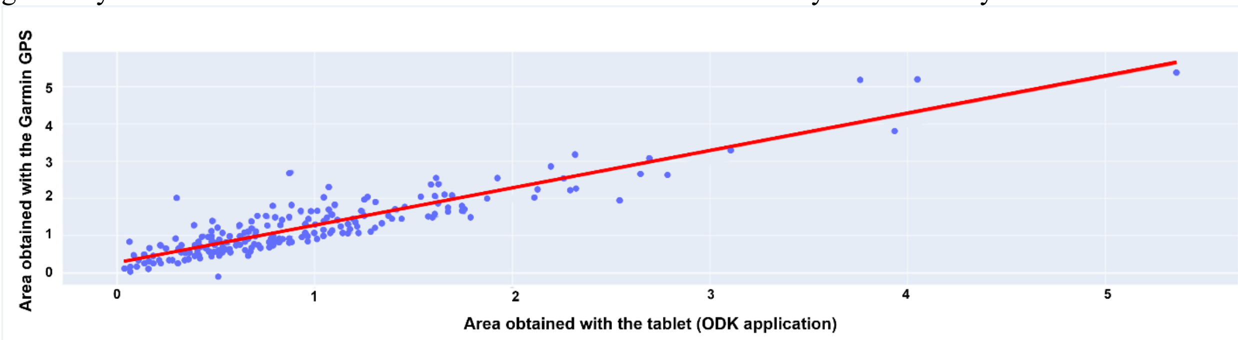


Figure 8. Predicted areas by the ODK app on an Android tab (blue points), the red line is the regression line related to the predicted dataset and the green line, the reference related to the GPS dataset

Several measurements deviate significantly from the reference, which comes from the use of the manual mode in ODK (i.e. enumerators to manually encode GPS measures at the inflection points of the plot). Errors in the area measurements were caused by:

- errors in encoding the points by the enumerators;
- waiting time: the tablet's GPS takes time to stabilize its position with acceptable accuracy. A shortened waiting time can then lead to an error of up to 30 meters from the position of the investigator;
- recording errors by the tablet: it is possible that the device or application does not record a point correctly or does not record it at all.

Statistically and visually, the tablet dataset is less accurate than the one recorded with the GPS and has a bias that underestimates the baseline value. The situation is improving when using ODK in automated mode (Figure 9), but the comparative analysis did not support the claim that a tablet can replace a GPS for area measurement.

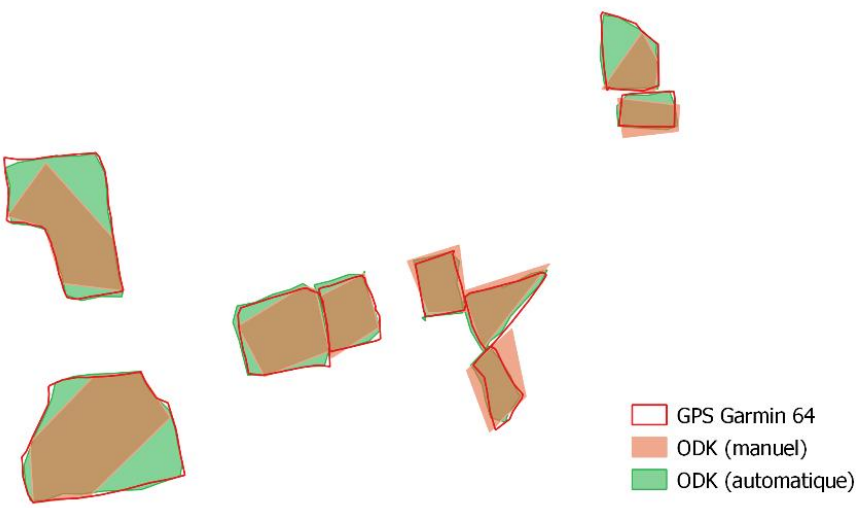


Figure 9. Examples of parcel's delineations by the GPS in red, by the tablet with ODK Collect in manual mode in orange and in automated mode in green

Additional non-crop class data

As part of the pilot, enumerators were instructed to collect data on non-crop land cover classes (e.g., fallow, built-up, water) through the ODK application. However, only a limited number of non-crop polygons were recorded. Several key classes were missing. To fill this gap, additional non-crop features were manually delineated using very high-resolution satellite imagery. In total, 50 non-crop polygons were available for model calibration.

Crop Classification and area indicators

Sen4Stat solution

The crop type classification was done using the Sen4Stat system (European Space Agency, 2019), an evolution of the Sen2Agri system, whose performance has been demonstrated and validated in a variety of contexts (Defourny et al., 2019).

The Sen4Stat toolbox is a standalone operational processing chain which aims at facilitating the uptake of Earth Observation data in the official processes of National Statistical Offices, since the early stages of the agricultural surveys to the production of statistics. It automatically ingests and processes Sentinel-1 and Sentinel-2 time series in a seamless way for operational crop mapping and yield modelling, using ground data provided by national statistical surveys. It generates a set of EO-based products, such as time series of biophysical indicators, crop type maps and yield estimates. These EO products are then integrated with the agricultural survey to improve the statistics. Different types of statistics improvements are targeted by the system: (i) reduction of the estimates error, (ii) disaggregation to smaller administrative units, (iii) provision of timely crop area and yield estimators and (iv) optimization of sampling design by using maps to build or update sampling master frames. EO products are also used to support the quality control of the survey data.

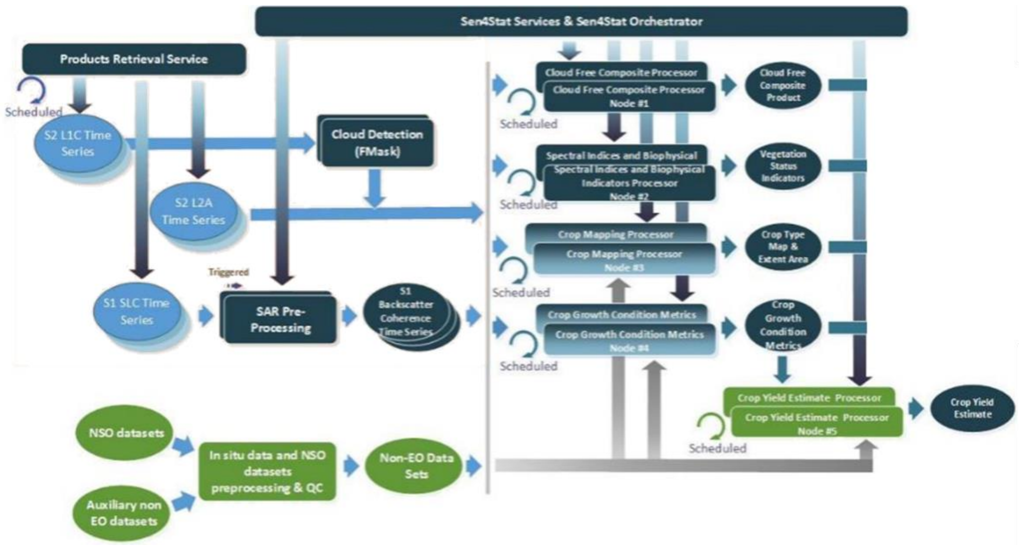
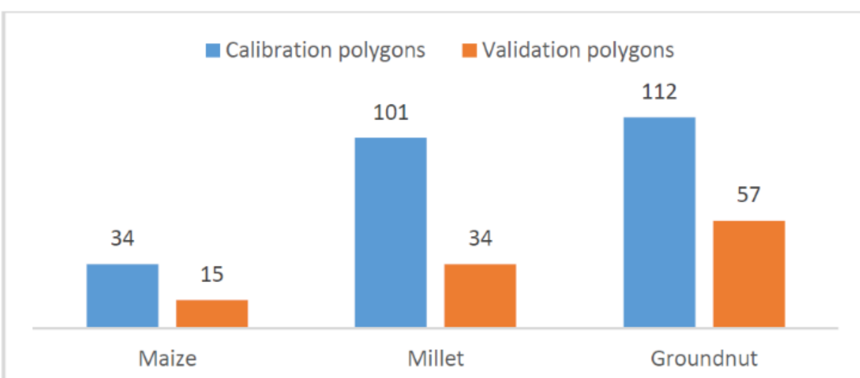


Figure 10. Open-source Sen4Stat toolbox

The system has been tested and demonstrated in various countries around the world, thus addressing a wide diversity of both cropping systems and agricultural data collection protocols. It is available for download on the Sen4Stat website (https://www.esa-sen4stat.org/).

Nioro crop type map 2021

The crop type map was generated at the end of the season from May 1, 2021 to December 31, 2021, using Sentinel-2 time series. The training dataset used contains the 50 non-crop polygons and the 247 crop polygons obtained after joining and cleaning the ODK and GPS dataset (Figure 11 – calibration polygons). The distribution of observations is very uneven for the different crops and insufficient for maize.



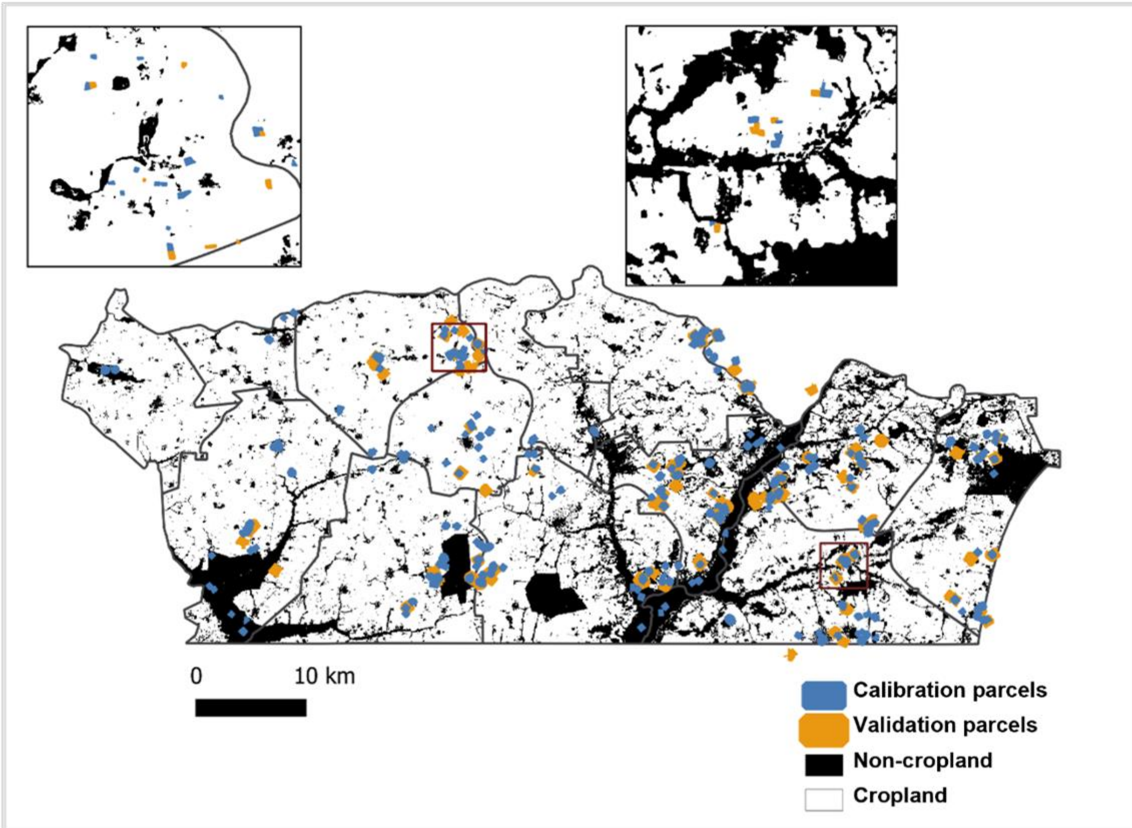


Figure 11. Distribution of in-situ data into a calibration data set and a validation data set

The obtained crop type map is shown in Figure 12.

The confusion matrix (Table 1) indicates an overall accuracy of 97.1% for the cropland mask and 88.2% for the crop type map. F-score values were 98% for cropland and 95% for non-cropland. For specific crops, F-scores reached 95.2% for ground nuts, 83.8% for millet, but only 54.8% for maize, which showed substantial omission. The representation of the three crops in the dataset was highly unbalanced, and the number of samples is clearly too low to allow a proper maize classification.

Table 1: Confusion matrix (expressed in number of pixels) for the crop type map, with contamination and omission values for each crop, UA as user accuracy and PA as producer accuracy

Expressed in number of pixels		Field survey: Non crop	Maize	Millet	Groundnut	UA	Contaminations (%)	Omissions (%)
Crop type map	Non crop	2205	0	34	17	97.7	2.3	9.3
	Maize	0	325	10	0	97.0	3.0	47.9
	Millet	202	265	2755	3268	80.5	19.5	12.6
	Groundnut	25	34	354	3487	88.8	11.2	6.3
	PA	90.7	52.1	87.4	93.7			

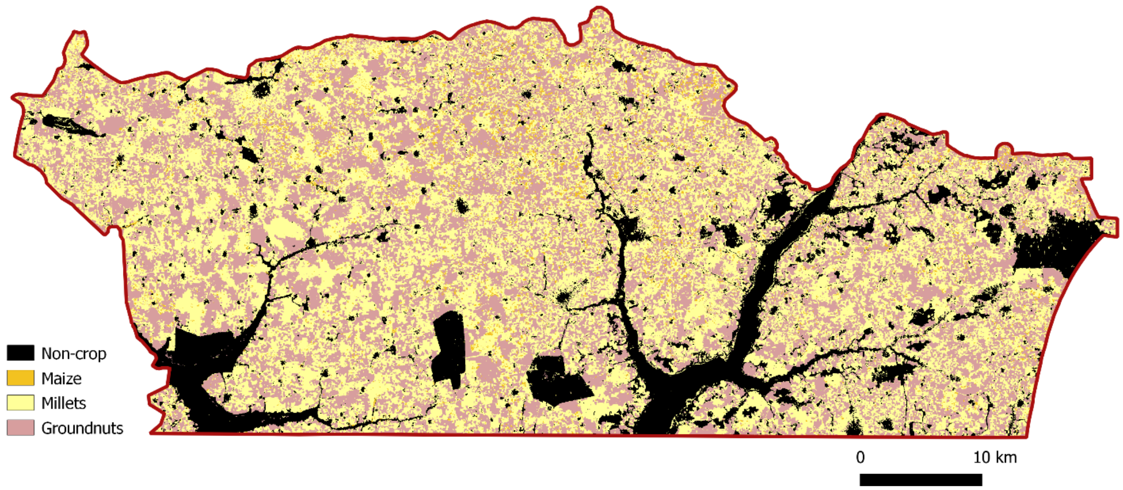


Figure 12. Crop type classification of the Nioro department based on satellite imagery from May 1, 2021 to December 31, 2021

Acreage estimation by integrating survey and EO data

Producing statistics by counting those pixels classified as a given crop type does not provide unbiased area estimates. Maps are subject to omission and commission errors (Czaplewski, 1992), which are linked to the ability of the classification method to distinguish between classes. Regression estimators and calibration estimators have been widely used to correct such biases (F. J. Gallego and Rueda, 1993; Khan et al., 2018; Olofsson et al., 2014; Li et al.,2023). They are traditional ways to combine accurate, possibly unbiased, information, observed on a sample with less accurate, biased information, known for the whole population or for a larger sample. Said differently, maps are used to improve an estimator that has been computed from a ground survey on a sample in preserving as much as possible the properties of the ground survey estimators (unbiasedness) and reducing the variance (Gallego 2004).

Because the sampling design in Senegal is a list frame, the integration of remote sensing and ground data cannot be done using linear models, but with multinomial logit models. These multinomial models deal with the uncertainties and generate probabilities that a pixel of a given class in the map is actually this given crop on the ground. Despite the fact that the map was not perfect, this model successfully demonstrates the added-value of EO data for agricultural statistics estimation.

Figure 13 presents the crop acreage estimates in the department of Nioro, for the two main crops which are millet and groundnut and shows the efficiency of using the crop type map to support this estimation of crop acreages.

Crop type	Acreage (hectare)	Uncertainty				
		Standard error	Coefficient of variation (%)	Limits of 95% confidence interval		
				Lower	Upper	Amplitude
Millet	89215	3661.103	4.11	81978.88	96330.4	14351.52
Groundnut	78815	2923.94	3.71	73089.15	84550.98	11461.82

Crop type	Standard errors of proportion estimators		Relative efficiency of Classified RS data
	Using only ground data	Using ground & RS data	
Millet	3.37	1.90	3.13
Groundnut	3.34	1.52	4.80

Figure 13. Crop acreage estimates using EO and ground data in Nioro and efficiency of using the crop type map for crop acreage estimation (bottom)

The use of a wall-to-wall map also allows to make acreage estimates available at the "arrondissement" levels with a reasonable error (expressed as the coefficient of variation), as shown in Figure 14.

Arrondissement	Millet		Groundnut	
	Acreage (has.)	Error (CV%)	Acreage (has.)	Error (CV%)
Medina Sabakh	20067.21	8.6	19765.36	7.3
Paoskoto	38316.02	5.3	35018.93	4.0
Wack Ngouna	30831.77	11.9	24030.71	10.7
Total Nioro	89215,00	4.11	78815,00	3.71

Figure 14. Crop acreage estimates at the district (arrondissement) level

3.3. Lessons learned

The alignment of the AAS protocol with FAO's in-situ data quality framework significantly enhanced the compatibility of field data with EO applications. Classification accuracy improved from below 78% (with the original 2018 AAS data) to over 85% following the pilot implementation in 2021–2022. This result has to be confirmed by an implementation at larger scale, but it already nicely illustrates the impact of structured, high-quality, plot-level georeferencing.

Moreover, integrating satellite imagery with field data enabled the use of model-assisted estimators—such as regression estimators—for crop area statistics. This approach yielded more accurate results, with lower coefficients of variation, compared to using field data alone.

Building on these outcomes, DAPSA endorsed the revised protocol and extended its implementation to six additional districts in 2024. This successful institutionalization underscores the value of diagnostic frameworks, phased piloting, and continuous quality improvement in advancing the operational use of EO in national agricultural statistics.

4. **Case Study 2: Uganda**

Between 2019 and 2021, FAO and UCLouvain collaborated with the **Uganda Bureau of Statistics (UBOS)** to pilot EO-compatible data collection protocols in two districts—**Lira** and **Tororo**—within the agro-ecological zone managed by the **Zonal Agricultural Research and Development Institute (ZARDI) of Bulindi**. The pilot region corresponded to a single agro-ecological unit, providing a coherent zone for remote sensing calibration and statistical analysis.

The agricultural survey in place in Uganda is a list frame survey over all country. In 2019, no GPS points were collected at the parcel-level. The initial plan was to work hand-by-hand with UBOS to test an adjustment of their protocol in the pilot region to make it EO-compatible. Unfortunately, with the COVID lockdowns in 2019-2020, survey activities were suspended, and the project had to design and implement an ad-hoc survey completely independent from the UBOS official survey.

Sampling Design and Data Collection

A total of 44 Enumeration Areas (EAs) were selected across the two districts (Figure 15). Within each EO, a two-stage area sampling frame was adopted, consisting of 9×9 clusters of 9 observation points each, spaced approximately 200 meters apart (Figure 16). This design was intended to optimize spatial coverage and ensure alignment with the resolution of medium-resolution EO imagery, such as Sentinel-2, while maintaining operational feasibility.

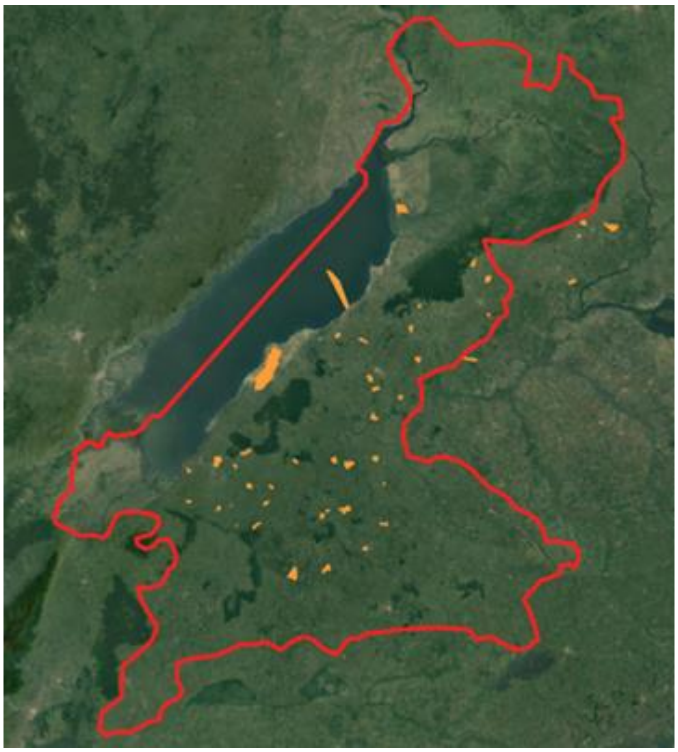


Figure 15. Enumeration Areas from the AAS survey included in the Bulindi ZARDI (in orange)

A protocol was designed, aiming at collecting crop type information and GPS coordinates at the parcel-level. The field campaign was planned to last around 30 days working with two teams of 5 enumerators + 1 driver each.

The protocol for each EA was as follows (Figure 16):

- The survey of one EA was designed to last around one half day;
- One EA was made of 9 clusters of 9 points each;
- Each team was split in 2 pairs of 2 enumerators and it was proposed that each pair starts the assessment at opposed corners of the EA:
    - Each team visited 4 clusters;
    - The first team that finished its 4 clusters was assessing the fifth/middle cluster.

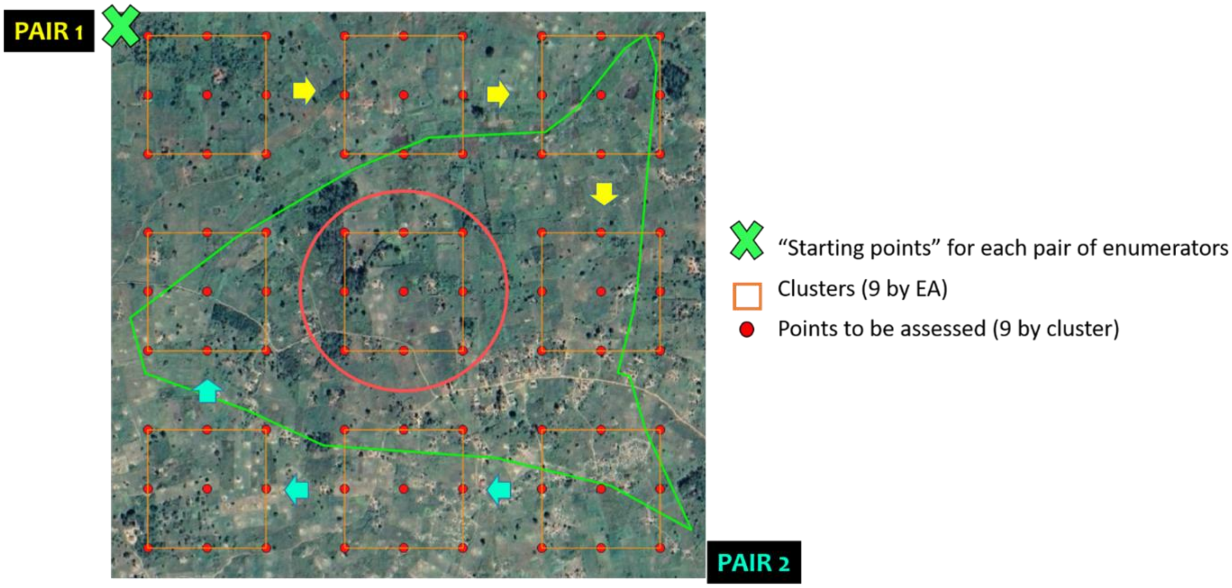


Figure 16. 9 clusters of 9 sampling points within one EA and protocol to survey one EA, made of 9 clusters of 9 points

The protocol relied on the App ODK Collect on tablets and Garmin eTrex Touch 35. It was planned for enumerators to upload the filled-in questionnaire (with ad-hoc GPS records) every night. This was perfectly done for the questionnaires, but not for the GPS records because of deficient material (no cable to link the GPS to a computer).

Preliminary Findings and Data Quality Challenges

The geospatial sampling design ensured statistically robust spatial coverage. However, challenges emerged during data collection:

- Crop misclassification was observed in several cases due to difficulties in crop identification at early growth stages.
- Some enumerators struggled to distinguish between mixed cropping systems or intercropped fields.
- While the spatial layout and GPS coordinates were precise, several entries lacked complete metadata or displayed inconsistencies in attribute labels.

The FAO in-situ data quality framework was partly implemented for this pilot, including spatial referencing and basic metadata capture, no post-hoc spectral consistency checks (e.g. NDVI time-series validation, buffer analysis) were applied.

EO-Based Outputs and Initial Results

The EO analysis component of the pilot involved processing Sentinel-2 time series imagery over the Bulindi AOI. Using a supervised classification approach, crop type maps were generated (Figure 17) and validated against field points (Table 9). The classification covered key crops including maize, cassava, ground nuts, and millet.

- Initial accuracy assessments using a subset of 1,980 field observations indicated an overall classification accuracy of 78.4%, with F1 scores of 86% for maize, 75% for groundnuts, and 72% for cassava.

- Classification errors were mainly observed between mixed fields and fallow/grasslands, underscoring the need for post-stratification or farmer-reported validation in future protocols.

Table 9. Confusion matrix of the Crop Type map with pure crops only

		Map				
		Maize	Cassava	Sweet potato	Sugarcane	
Reference	Maize	320	21	0	4	93%
	Cassava	32	38	0	3	52%
	Sweet potato	10	9	0	2	0%
	Sugarcane	12	13	0	44	64%
		86%	47%	NA	83%	79%

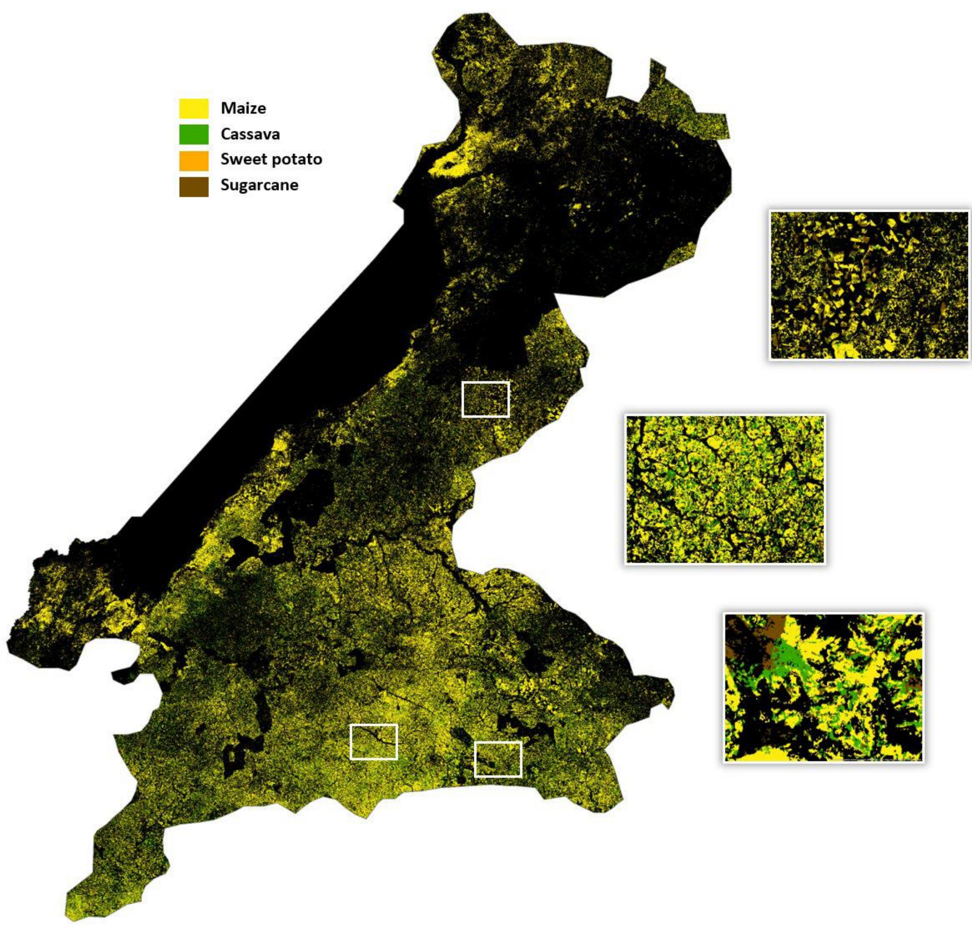


Figure 17. Overview of the annual crop type map focusing on pure crops at national scale, with the non-cropland areas masked in black

EO-Derived Area Indicators

In addition to crop type classification, the Uganda pilot explored the use of satellite-derived crop area indicators to provide early estimates of cropland extent and crop-specific distribution. These indicators are not intended to replace official agricultural statistics but rather to offer timely insights months before final survey results become available.

Area estimates were computed through pixel counting using the classified Sentinel-2 crop mask and crop type maps. Within the Bulindi ZARDI agro-ecological zone, the total cropland area was estimated at 7,795 ha, representing approximately 26.6% of the total area, with the remaining 21,462 ha (73.4%) classified as non-cropland.

The four main crop types identified included maize, cassava, sweet potato, and sugarcane, with maize being the dominant crop as shown in Table 10.

Table 10. Pixel counts of the main crop types at the Bulindi ZARDI level

	Number of pixels	Hectares
Maize	50.986.156	5,099 ha
Cassava	19.555.038	1,956 ha
Sugarcane	7.286.376	728,64 ha
Sweet Potato	127.244	12,72 ha

However, pixel-based area estimates are inherently affected by classification errors, particularly commission (false positives) and omission (false negatives). Corrections were applied using class-specific error rates derived from the confusion matrix (Table 11):

Table 11. Omission and commission errors associated with the four main pure crops (derived from Table 3-1) and impact on the area estimators calculation

	Omission error	Commission error	
Maize	7%	14%	Commission > omission => decrease the estimates by ~7%
Cassava	48%	53%	Commission > omission => decrease the estimates by ~5%
Sweet potato	100%	N/A	
Sugarcane	21%	17%	Omission > commission => increase the estimates by ~4%

Despite these limitations, the EO-derived area indicators proved valuable in demonstrating how early estimates of production zones can be generated at the subnational level. This approach could be particularly useful for planning surveys, allocating resources, or triggering early warning mechanisms in years of climatic stress.

Efforts to compare EO-derived estimates with official statistics were constrained by the lack of disaggregated ground truth data and the absence of recent FAOSTAT records for 2020–2021 at subnational level.

Lessons Learned and Recommendations

These preliminary results underscore the strategic value of harmonizing survey protocols with EO classification workflows. While the pixel-based estimation approach remains constrained by classification

uncertainty, its timely outputs can complement conventional survey statistics, especially in contexts where delays, limited coverage, or budgetary constraints inhibit comprehensive field data collection.
The Uganda case highlights the importance of not only strengthening the spatial design of field surveys but also ensuring that validation protocols—such as the integration of confusion matrices and error correction—are systematically applied. Embedding these checks within national workflows will be essential to transition from demonstration pilots to institutionalized, quality-assured EO-based agricultural statistics.

The Uganda pilot produced actionable insights for scaling EO-compatible methods across the national statistical system. A 2024 feasibility study led by FAO and UBOS assessed options for integrating EO into the Uganda Household Integrated Survey (UHIS). Three options emerged:

1. Farmer-assisted parcel geolocation, where respondents identify their plots using satellite base maps.

2. Enumerator-based polygon mapping with GPS-enabled tablets to capture field boundaries and link them to survey records.

3. Windshield surveys, where trained enumerators rapidly assess visible crops from vehicles, suitable for stratification or post-classification refinement.

Together, these strategies offer a scalable pathway for mainstreaming EO into official agricultural statistics. The Uganda case highlights that even with limited in-situ controls, EO can provide statistically meaningful outputs—if quality assurance protocols are institutionalized, and training is sustained.

5. Conclusions and Recommendations

3.1. What the two cases jointly show

The Senegal and Uganda case studies demonstrate, from two different starting points, that the bottleneck to scaling EO for official agricultural statistics is rarely the satellite data or the algorithms—it is the fitness-for-use of in-situ data and the absence of systematic quality assurance protocols.

- Senegal shows what happens when the FAO in-situ quality framework is applied end-to-end: plot-level georeferencing, daily feedback loops, and EO-based post-hoc validation lifted overall classification accuracy from ~78% to >85% and enabled more precise, model-assisted area estimation. The protocol has since been institutionalized and scaled to six districts (2024).
- Uganda shows the other side of the coin: a solid spatial design and georeferencing, but gaps in semantic accuracy and post-hoc EO validation (e.g., no NDVI/spectral purity checks). The result was a workable (≈78% OA) but less reliable classification and area indicators that required heavy error corrections. The 2024 feasibility study therefore focused on how to retrofit EO-readiness into national survey operations (UHIS).

Together, the two experiences validate FAO's two-pillar framework (survey design + EO-based post-hoc assessment) as a practical diagnostic and reform tool, and they clarify the minimum conditions NSOs must meet to operationalize EO for statistics.

3.2. Cross-cutting conclusions

1. *Survey protocols must be (re)designed for EO*
   Plot-level polygons (not single GPS points), local homogeneity aligned with the satellite footprint, crop observability at the right phenological window, balanced samples for minority crops, and rich contextual metadata are non-negotiable.

2. *Quality assurance is a process, not a checkbox*
   Real-time feedback loops (as in Senegal), systematic post-hoc EO validation (NDVI time-series clustering, spectral purity, topology checks), and confusion-matrix based corrections must be institutionalized—not left to projects.

3. *Semantic accuracy matters as much as spatial accuracy*
   Uganda shows that even when points are well georeferenced, label noise (crop misidentification, mixed cropping ambiguity) degrades model performance and downstream statistics.

4. *Rebalancing and representativeness are critical for ML training*
   Class imbalance (e.g., over-represented groundnut/rice in Senegal; dominant maize in Uganda) must be actively managed to avoid biased classifiers.

5. *EO can already deliver policy-useful "early indicators"*
   Pixel-count–based area indicators (bias-corrected via the confusion matrix) can provide timely, sub-national signals—useful for planning, early warning and targeting—but only if transparently qualified and accompanied by uncertainty metrics.

6. *Institutionalization requires phased testing + capacity + governance*
   Senegal's uptake shows that pilots → protocol refinement → formal endorsement → scale-up is a workable pathway. Uganda's feasibility study highlights the need to embed EO-readiness into national survey workflows (e.g., UHIS) from the outset.

7. *Technology is available; process and ownership are the gaps*
   Open solutions like Sen2-Agri / Sen4Stat, cloud platforms, and low-cost mobile tools (ODK, Survey123) are mature enough. The binding constraint is consistent funding for training, QA, and national stewardship.

3.3. Recommendations to AFCAS Members

Based on the quality framework and two case studies presented in this paper, AFCAS are invited to take note of the following recommendations and best practices:

A. <u>Policy & Governance</u>
1. **Endorse and adopt FAO's in-situ data quality framework** as the regional reference for EO-readiness of agricultural surveys and censuses.
2. **Mandate national EO–statistics task forces** (NSOs, Ministries of Agriculture, Space/RS agencies, universities) to steer protocol revisions, QA processes, and scale-up plans.

B. <u>Methods & Standards</u>
1. **Embed the two-pillar QA workflow** (design diagnostics + EO post-hoc validation) in national survey SOPs, with clear acceptance/rejection criteria for datasets.
2. **Standardize core metadata** (e.g., plot geometry, crop phenology stage, management practices, heterogeneity flags) to enable interoperability across surveys and EO pipelines.
3. **Require confusion-matrix–based bias correction** (and uncertainty reporting) for all EO-derived area indicators used in official or quasi-official reporting.

C. Operations & Pilots

1. Run **phased pilots** before nationwide roll-out (Senegal model): start in 1–2 districts, iterate protocols with daily QC, then scale.
2. Prioritize farmer-assisted geolocation, enumerator polygon capture, or windshield surveys (Uganda options) when full plot delineation is not yet feasible.

D. Capacity Development

1. **Invest in applied, on-the-job training** on:
   a. EO-based QA (NDVI clustering, spectral purity checks, topology validation)
   b. Class rebalancing and sampling for ML
   c. Model-assisted estimation (e.g., regression estimators) for crop areas
2. **Create a regional peer-learning loop** (e.g., DAPSA–UBOS exchanges) to accelerate transfer of working protocols, scripts, and dashboards.

E. Infrastructure & Sustainability

1. **Leverage shared cloud platforms** (e.g., UN Global Platform) for scalable, secure pre-processing and versioned workflows, with a cost model aligned to national budgets.
2. **Plan for maintenance**: dedicate resources for routine QA, annual protocol reviews, and periodic external audits of EO-statistical pipelines.

3.4. Closing message to AFCAS Members

If AFCAS members want EO to move from project pilots to production statistics, survey protocols, QA routines, and institutional mandates must be modernized together. The Senegal and Uganda experiences show that where frameworks guide practice, accuracy, efficiency, and uptake follow; where they don't, EO outputs remain ad hoc and hard to trust.

AFCAS can catalyze this transition by considering the adoption of FAO framework, backing national pilots with targeted capacity building, and monitoring country progress against a small set of EO-readiness indicators.

References

1. de Simone L, Pizarro E, Paredes J, et al. Quality control of training samples for agricultural statistics using Earth observation. *Statistical Journal of the IAOS*. 2025;41(2):467-481. doi:10.1177/18747655251338033
2. EOSTAT FAO Technical Reports: Senegal (2021, 2022); Uganda (2021); Uganda Feasibility Study (2024).