







Data & Digital Solutions

Ghana Cropland Mapping - a VIFAA Case Study

Data-Driven Responses to Africa's Fertilizer Needs

Fertilizer use is one of the most important drivers of agricultural productivity. Despite this importance, fertilizer use across many countries in sub-Saharan Africa is far below the global average or government recommendation. While there is interest among the public and private sectors to increase fertilizer use and supply, there is a widespread lack of basic fertilizer consumption data at the national scale.

To address this issue, the Bill and Melinda Gates Foundation provided funding for **Development Gateway: An IREX Venture** (DG) and AfricaFertilizer (AFO) to implement the Visualizing Insights for Fertilizer in African Agriculture (VIFAA) program to systematically collect this data and make it publicly available in order to address data shortages relating to total crop production. As part of VIFAA's work **Quantitative Engineering Design** (QED), which has a track record of mapping land cover, was selected to use satellite imagery and artificial intelligence (AI) to map all croplands and calculate cropland estimates across Ghana; this work builds upon the cropland mapping of Nigeria in 2020. Through the VIFAA Program, DG and partners will use the resulting data and maps to answer two questions: What is the total cropland under production in Ghana (to be answered during phase 1) and what is the cropland under production, by crop type (to be answered during phase 2).

Under the VIFAA program, we brought together the cropland map, plant directories, and fertilizer retail prices to offer geospatial intelligence that is invaluable in enhancing agricultural productivity, managing crops, monitoring land use changes, and assessing the environmental impact of farming activities. This data is beneficial to both private and public sectors as they plan and execute agricultural expansion programs.

By integrating these datasets, agricultural practitioners gain a better understanding of the resources available in their local area and their prices, empowering them to make informed decisions on crop selection and resource management. The maps also assist in identifying potential opportunities, such as nearby transportation routes.

Definitions

1. Croplands: QED defines croplands to include land under temporary crops (within season double-cropped areas

are counted only once); land under market and kitchen gardens; permanent crops; and land cultivated with crops that occupy the land for long periods and need not be replanted after each harvest, including flowering shrubs, fruit trees, nut trees, and vines. Excluded are pasture lands (both temporary and permanent), fallowlands, wood, and timber trees. Consequently, QED's definition only encompasses land that is already cultivated which includes all crops that can use fertilizer, including permanent and perennial crops and excludes pasture lands.

2. Agroforestry: In Ghana, QED differentiated between agroforestry and non-agroforestry croplands. Tree crops were classified based on visual identity patterns associated with agroforestry systems as a subset of general cropland. Non-agroforestry croplands were areas classified as croplands (using the definition outlined above), which did not have visually identifiable agroforestry.

Mapping Methodology

QED's high level approach to mapping croplands and building structures is four-fold:

- 1. **Surveying:** Generate training data by labeling satellite imagery and looking at the visual patterns in the landscape between different land uses.
- 2. Modeling: Build multiple AI models using the training data so that the computer learns from the training data to classify unknown areas across the country.
- 3. Validation:

(a) *Mapping*: Generate predictive maps with each model, and assess each one's performance.(b) *Refinement*: Repeat steps 1-3 in selected areas requiring further improvement.

4. Statistics and Visualization: Generate summary statistics based on predictive maps.

Surveying: The quality of AI outputs are strongly dependent on the quality of the training data. Training AI requires providing many example classifications to a computer; these types of data are referred to as training data. To meet this training need, QED developed Geosurvey, a software tool for efficiently collecting and labeling visual training data. This tool feeds labeled images of agricultural landscapes into a computer program, where trained surveyors assess each image and classify the features within each image (i.e., croplands or not). The process requires a team of skilled surveyors to identify visual indicators of agricultural features and learn patterns which can change depending on the cropping system and geographic region where the specific training data was developed to match the unique locations.

The classification work done by the trained surveyors is also closely supervised by QED's in-house agronomist and further reviewed by an administrative panel for structured annotation and bi-directional communication (whereby surveyors and agronomists can continue dialogue). Due to the variety of agro-ecological zones (AEZ) in Ghana, the landscape data fed into Geosurvey was broken into agro-ecological zone-specific modules.

As tree crops (principally cocoa, cashew, and oil palm) are critical to Ghana's agricultural industry and fertilizer sector, QED separated the classification for agroforestry and nonagroforestry croplands. To do this, they created separate markers within Geosurvey for surveyors to differentiate between these two categories. Surveyors outlined these regions as a separate category of training data. This enabled them to produce models specifically for agroforestry.



Ground Data Collection Campaigns

QED partnered with two organizations, Farmerline and African Fertilizer and Agribusiness Partnership (AFAP), to conduct ground data collection campaigns in Ghana. These campaigns produced over 7,000 data points randomly distributed across all of Ghana. The ground data is not used to generate models, but rather to provide independent and contemporary assessment of the mapping outputs. This data was also used to compare the accuracy between QED's maps and other mapping products.

Supervised Learning and Active Learning

QED constructed machine learning models to infer patterns between satellite imagery and training datasets, and then extrapolated these patterns over the land areas of Ghana. Active Learning workflow, a modeling framework which evaluates regions where the models have lower confidence, was implemented. Additional datasets are generated from the regions, resulting in improved model performance while increasing efficiency for collecting targeted additional data.

Cropland Statistics

Statistics have been generated for cropland areas under production across the various regions in Ghana (see table below). These statistics include total land area, estimated cropland area, and estimated cropland percentage, which can be compared against statistics from other sources.

Region	Total Area (sq.km)	Estimated Cropland Area (sq.km)	Estimated Cropland percentage
Ashanti	24904.19	3230.14	12.97
Brong Ahafo	39488.97	5341.26	13.53
Central	9651.72	825.93	8.56
Eastern	18501.06	1163.09	6.29
Greater Accra	69468.08	1089.28	29.53
Northern	83,226	20888.61	30.07
Upper East	8603.85	3893.18	45.25
Upper West	18906.44	3913.68	20.70
Volta	20492.51	5263.28	25.68
Western	24589.04	155.57	0.63

Ghana Cropland stats - 2021

The QED maps had accuracies greater than 85% due to the region-specific models trained from data generated from within the AEZs. This helped increase accuracy by providing similar data to both the surveyors and the AI. Performance metrics for the cropland and agroforestry maps produced by QED for Ghana in 2021 are:

- Total cropland estimate for Ghana: 49,764.0 km2
- Agroforestry: 4,000.0 km2
- Non-Agroforestry Cropland: 45,764.0 km2



Ghana Cropland Mapping - a VIFAA Case Study



Ghana Cropland Map 2022 (Chloropleth map)

Limitations

High Cloud Coverage: Cloud coverage is a common challenge when analyzing land cover using satellite imagery. In fact, areas close to the equator (such as Ghana) are covered by clouds for most of the year, making cropland mapping very challenging. To address this, we developed a solution that takes advantage of Sentinel 19s satellite radar, which operates at wavelengths unimpeded by clouds. Using this, we extracted high-resolution time series data that helped us greatly enhance model performance.

Field Data Collection: Two field data collection campaigns were executed for Ghana — one with Farmerline and another with AFAP. Limited data was collected during the off/dry-season (approximately November 2021 - April 2022). This was due to limited time to work with ground teams from commercial partners like Farmerline and the need to prioritize data collection in the main crop production season (May - October 2021). In addition, despite extensive and persistent efforts to obtain field survey data to build better open access maps and allow independent evaluation, QED was not successful.

Next Steps & Recommendations

Off-Season Mapping: Collecting data on crop production during the off-season can greatly improve the accuracy and reliability of agricultural mapping and analysis. Although this field may not have been prioritized in the past due to the high reliability of independent validation from main production season data, progress has been made in modeling and integrating transfer learning with satellite data. By gathering more off-season data, we can gain a better understanding of crop growth and distribution patterns throughout the year, which can inform future crop management decisions and promote sustainable agricultural practices. This data can also be used to enhance the precision of satellite imagery and other mapping technologies, leading to more effective agricultural analysis. Ultimately, investing in off-season data collection can optimize crop yields and minimize waste.

Satellite Imagery Availability: Access to affordable, highresolution, national-scale satellite imagery remains a constant limitation for efforts in cropland mapping. However, we leveraged our ability to successfully develop transfer learning models seeded with higher resolution imagery from previous years. These models then served as the basis for developing contemporary models using lower-resolution imagery covariates. This approach can unlock insights into future mapping efforts. The challenges with achieving high recall are well-recognized, and to our knowledge, we have no prior tree plantation maps to compare against.

Cropland Under Production by Crop Type: Revisit the question around cropland under production by crop type. For this data to be accurately collected, we need continuous and widespread crop type data that is merged by multiple parties who can all benefit from the results. We hope to be able to build pathways between the private sector and the government to be able to pursue this in the near future.

Interested in learning more about the cropland mapping in Ghana and VIFAA? Reach out to Ousmane Kone at <u>okone@developmentgateway.org</u> and Sebastian Nduva at <u>snduva@ifdc.org</u>.