NOV

2025

Drone Mapping Report

Application of Drone Solutions for Sugarcane Mapping.

Western Catchment Region

Kenya Sugar Board

ECOSPACE SERVICES

Providing Geospatial Solutions

Table of Contents

1 Executive Summary

The Western Catchment drone mapping exercise was executed to modernize cane census and supply planning for the Kenya Sugar Board (KSB). The project combined high-resolution drone imagery, field digitization, and satellite-based analysis using Google Earth Engine (GEE) to estimate area under cane cultivation, and cane available aged 12 months and above. Drone imagery covering over 200 fields was mosaiced via DJI Terra and digitized in QGIS to produce farm polygons with attribute metadata. Vegetation indices (NDVI, EVI, SAVI) were computed in GEE for the digitized farms; those indices and polygons provided the labelled training dataset for supervised classification using Sentinel-2 time series.

Concurrently, ODK-driven conventional cane census recorded spatial locations and age-class data for cane 12 months; these ground records were digitized and used to train GEE models to identify spectrally similar mature-cane patches. Analysis revealed substantial overestimation in the conventional census: satellite and drone-based estimates indicate relatively low value compared to census to sustain continuous mill crushing. The exercise provided substantial operational training for KSB staff in drone operations, data processing, and GEE workflows. Key constraints experienced were fragmented small fields from land subdivision, weather/cloud, limited battery capacity, storage, transport logistics, and unstable power have been documented with targeted recommendations. We request mill spatial datasets (crop cycles, planting dates, varieties) and enhancements to the ODK instrument to improve integration. This report details methodology, results, accuracy assessment, implications for milling, and recommendations for future operationalization.

2 Background, Rationale and Literature Context

Accurate area and age-class estimates of sugarcane are vital for mill scheduling, supply contracting, and national industry stability. Traditional field-based cane censuses are labour-intensive, limited in spatial coverage, and prone to sampling and reporting bias. Remote sensing offers scalable, objective alternatives i.e. combining high-resolution drone imagery for local accuracy and Sentinel-class satellites for region-wide mapping. Studies from Brazil, India, and South Africa have demonstrated improved area estimates and yield proxies when combining UAV and satellite data with supervised classification and multi-temporal vegetation indices. For instance, Brazil's sugarcane monitoring programs use Landsat/Sentinel time series with random forest classifiers to map crop phenology and area; India's operational sugarcane studies apply Sentinel-2 NDVI/EVI to distinguish cane stages; South Africa has used UAV-derived canopy metrics to calibrate satellite-based production models.

Google Earth Engine (GEE) has become a preferred analysis platform for these applications due to its scalable processing, time-series handling, and ready access to Sentinel imagery. Drone platforms (e.g., DJI) provide centimeter to decimeter resolution, data essential for precise field boundary extraction and training data creation. Integrating ODK-grounded census points with UAV and satellite-derived spectral signatures increases classification reliability and provides an independent validation dataset. This combined approach addresses the key limitations of standalone census activities i.e. coverage, objectivity, and temporal consistency while delivering operationally actionable outputs for mill decision-making.

3 Study Area Description and Operational Context

The Western Catchment comprises heterogeneous landscapes spanning lower and upper catchment zones, dominated by smallholder and medium-scale farms. Climate is tropical to sub-tropical with bimodal rainfall; topography is undulating with pockets of low relief plains where sugarcane is concentrated. The region is characterized by increasing land subdivision, resulting in fragmented fields and variable field sizes. Farming systems include mixed cropping; sugarcane is interspersed with seasonal crops and fallow plots, complicating spectral differentiation.

Accessibility constraints (road condition, vehicle clearance) and weather variability, frequent cloud cover and rapid weather changes directly impacted field mapping windows. The Western Catchment's mosaic of small, sometimes narrow parcels necessitates high-resolution mapping for accurate area estimation. Drone operations were focused on representative zones covering over 200 fields stratified by sub-catchment, cropping pattern, and proximity to mills. Ground-truthing via ODK collected field level age-class data, planting dates where available, and farmer reported variety and management notes.

4 Data Collection: Drone Missions, Sensors and Ground Truth

Drone data were acquired using DJI platforms with multispectral/RGB sensors. Flight planning prioritized 70–80% front overlap and 60–70% side overlap to ensure robust mosaicking and orthomosaic quality. Typical flight altitude ranged 80–120 m AGL, delivering ground sample distances (GSD) of 3–5 cm for RGB missions which was sufficient for precise boundary delineation.

DJI Terra was used for mission planning and mosaicking, leveraging onboard geotags and GCPs where deployed. Field digitization in QGIS was performed from orthomosaics, producing farm boundary polygons with attributes (farmer ID, crop age, crop cycle, expected yield, drone area, last yield). Ground truth data were collected with Trimble Ecofield Tool, capturing GPS-located records of cane presence, age class (notably identification of fields older than or equal to 12 months), and management notes; these geopoints linked to digitized polygons.

Sentinel-2 Level-2A (surface reflectance) imagery provided multi-spectral temporal coverage; cloud masking and gap-filling strategies were implemented in GEE. Data management followed a structured pipeline: ingest drone Orthomosaic and digitized shapefiles into a repository, compute vegetation indices in GEE, extract index statistics per polygon, and assemble training/validation datasets. Data quality control included field inspections for GNSS error, orthomosaic seam checks, and visual verification of digitized boundaries. Storage constraints necessitated staged upload schedules; field backups were made on portable drives to mitigate single-point failure.

5 Preprocessing: Mosaicking, Georeferencing, and Polygon Extraction

Drone imagery mosaicking was completed in DJI Terra using tie-point matching and bundle adjustment to produce georeferenced orthomosaics. Where available, surveyed Ground Control Points (GCPs) were used to refine absolute positional accuracy; average positional error was quantified through RMS values per mission. Orthomosaics were clipped, corrected for lens distortion and exported as geotiffs for QGIS digitization. In QGIS, a standardized digitization protocol was followed: visible field boundaries approximated by vegetation change, fence lines, or farmer-provided markers were traced to create polygons.

Polygons were attributed with metadata: field ID, farmer name, area (computed), observed crop type, and estimated planting date if known. Edge-buffering rules and minimum polygon area thresholds were enforced to reduce digitization noise from narrow border features. Polygon geometries were simplified for computational efficiency while retaining fidelity for small-field delineation. For integration with GEE, polygons were exported as GeoJSON/GeoPackage and uploaded to an asset bucket.

Cloud masking for Sentinel-2 was performed using the scene classification layer and simple heuristics to remove high-cloud pixels. Temporal compositing strategies; median and 90th percentile composites, were tested to reduce residual cloud/noise. These preprocessing steps established a robust base for per-polygon spectral extraction and supervised classification.

6 Vegetation Indices, Spectral Metrics and Rationale

Vegetation indices (VIs) serve as proxies for canopy health, biomass, and chlorophyll content properties relevant to sugarcane detection and age-stage differentiation. Additional spectral metrics included Red Edge indices (available from Sentinel-2) to capture structural/morphological differences between cane and other crops and simple band ratios (e.g., NIR/Red) to enhance class separability. For maturity estimation (12+ months), temporal VI trajectories were analysed; mature cane shows distinctive seasonal stability and higher baseline NDVI/EVI values compared to annual crops. Per-polygon VI metrics were extracted i.e. mean, median, standard deviation, percentile values, and temporal slope, forming feature vectors for classifier training.

In GEE, VI computations were automated across entire time-series stacks and statistics were exported per polygon. The combination of drone-derived high-resolution texture and canopy structure (where multispectral drone data available) with satellite multi-temporal indices improved discrimination, particularly in fragmented small-field contexts.

7 Training Dataset Design and Sampling Strategy

A robust training dataset is essential for supervised classification. The drone-digitized farms provided high-fidelity labels for cane vs. non-cane classes and for age-state labels (particularly for fields greater than 12 months). ODK census points with age-class annotations were spatially linked to polygons to corroborate age labels. Sampling prioritized representativeness: stratified by sub-catchment (upper, lower), field size classes, cropping systems, and variety. Class balance techniques mitigated skew from dominant non-cane land cover.

For mature cane identification, training samples specifically targeted fields confirmed by normal census of over 12 months; spectral time-series for these reference polygons were collated to capture phenological signatures. Data augmentation included inclusion of multi-season Sentinel-2 composites to capture inter-annual variability and inclusion of different illumination conditions. An independent validation set (20–30% of labeled polygons) was reserved for accuracy assessment, ensuring no overlap with training polygons. Feature importance analysis informed dimensionality reduction: highly correlated indices were pruned, while texture measures from drone ortho (e.g., local variance) were retained when available. The final training set enabled implementation of machine-learning classifiers (random forest and support vector machines) within GEE, using cross-validation and confusion matrices to assess performance.

Classification was implemented in GEE, leveraging its large-scale temporal handling and built-in machine learning APIs. Two supervised classifiers were explored: Random Forest (RF) for robustness to noisy features and non-linearity, and Support Vector Machines (SVM) where margin-based separability proved advantageous. Input features comprised per-polygon VI statistics (NDVI, EVI, SAVI, red-edge indices), temporal slope values, band ratios, and, where applicable, drone-texture metrics. Pre-classification steps included standardization and outlier removal.

RF models used 100–200 trees with depth parameters tuned via grid search and k-fold cross-validation; SVM hyperparameters (kernel type, C, gamma) were tuned similarly. Class weighting addressed imbalanced datasets. Accuracy assessment utilized the reserved validation polygons generating confusion matrices, overall accuracy, Kappa statistic, producer's and user's accuracy per class. For area estimation, classified pixel counts were aggregated within the region mask and converted to area using pixel resolution scaling. Post-classification smoothing (majority-filter mode) addressed salt-and-pepper noise, with small isolated patches below a size threshold flagged for manual review. The classification output included probability/confidence rasters for cane presence and for mature-cane class to indicate model certainty and guide targeted field checks.

8 Results: Area Under Cane

a) The supervised classification, calibrated with drone-digitized polygons, produced a regional map of cane presence. Aggregated results indicate that approximate area under cane in the Western Catchment as at October 2025 is ..... lower than previously reported by the conventional census.

b) Key metrics: total mapped cane area = ...hectares (regional sum) distributed with ...% in upper catchment and ... % in lower catchment.

9 Results: Cane Available ≥12 Months.

c) Using ODK-identified mature-cane polygons as a training set, the model estimated area of cane aged 12 months and above. The GEE analysis yielded a mature-cane area estimate of **25,011.17** hectares. Spatial distribution indicates that mature cane is concentrated in discrete clusters rather than uniformly distributed which is important for mill logistics and haulage optimization.

d) When comparing mature-cane area to conventional census reports, the census tended to overstate mature-cane area by 16.99%, likely due to timing mismatches (post-harvest regrowth misclassified as mature), farmer reporting bias, or misinterpretation of age during field enumeration.

e) Importantly, the spectral signatures of mature cane exhibited stable high NDVI values over multi-month composites, a feature exploited by the classification. Temporal checks suggested that some fields marked as over 12 months had undergone recent ratoon or cutting events, explaining outliers.

f) The model confidence for mature-cane classification was lower than for general cane presence, reflecting greater spectral ambiguity; nevertheless, high-confidence mature-cane parcels were validated in the field.

g) Combining area estimates with average expected yields per hectare allowed an inference of potential tonnage available: this tonnage, cross-referenced with mill milling capacities and typical cut-to-crush efficiency, indicates an operational window of approximately four months continuous crushing before available mature cane would be depleted under standard throughput assumptions.

Area Under - Cane (12+ Months)		
No.	**County**	**Area (Ha)**
1	Bungoma	3,257.13
2	Busia	2,548.07
3	Kakamega	6,158.43
4	Nandi	4,272.19
5	Siaya	214.75
6	Trans Nzoia	3,165.97
7	Uasin Gishu	4,189.87
8	West Pokot	1,204.76
Total		**25,011.17**