A Sample-Based, Multi-Sensor Assessment of Land-Use and Land-Cover Change in Cameroon Using Collect Earth Online

Destin Loge Lokegna[1,2], Eloïse Guidi[3], Crystal Wespestad[4], Lucas Bembong[5], Robert Pismo[6], Maxime Ngemandji Moussa[6], Christophe Bring[6], Joseph Armathe Amougou[5], Timothée Kagonbe[6], Carine Bourgeois[1,2], Ashley Lehman[2], Rene Siwe Ngamabou[1,2]
1.SINREX Sustainable Solutions, Yaoundé, Cameroon; 2.Wilpa Capacity Development, Reston, VA, United States of America; 3. Coalition for Rainforest Nations, New York, NY, United States of America; 4. Spatial Informatics Group (SIG), Pleasanton, CA, United States of America; 5. National Observatory on Climate Change (ONACC), Yaoundé, Cameroon; 6. Ministry of the Environment, Protection of Nature and Sustainable Development (MINEPDED), Yaoundé, Cameroon

1–Context & Objectives

Cameroon's forests cover ~66% of the national territory (~31 Mha) and are central to climate regulation, biodiversity, livelihoods, and national commitments to climate, biodiversity, and land restoration under the Rio Conventions.

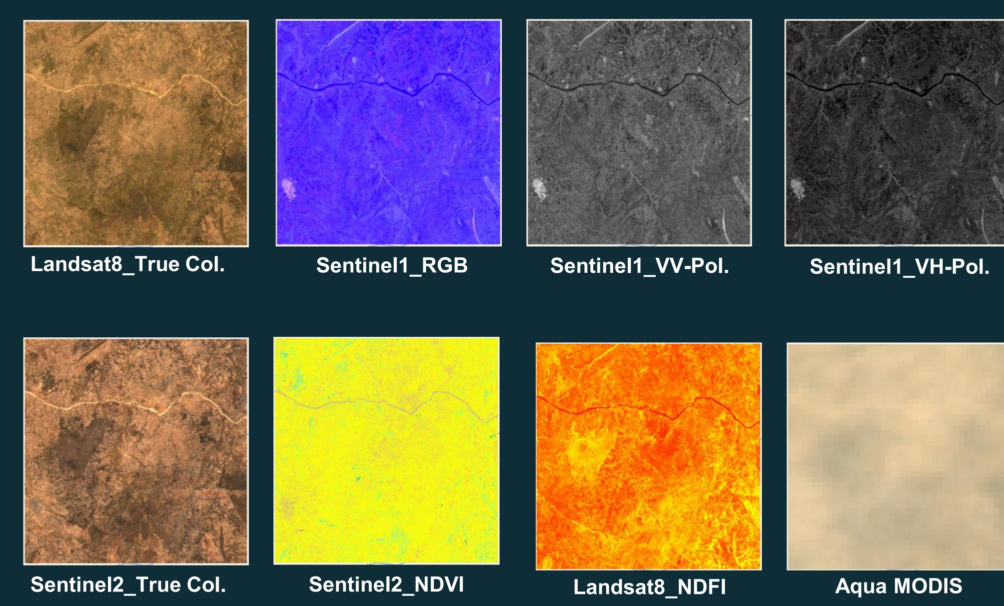
This work documents the national workflow implemented by the Ministry of the Environment, Protection of Nature and Sustainable Development (MINEPDED), and the National Observatory on Climate Change (ONACC) to produce IPCC-aligned estimates of land-use change and associated GHG emissions/removals (2000–2023) using Collect Earth Online (CEO).

Beyond producing estimates, the study formalizes a harmonized methodological framework combining sample-based interpretation, decision rules, and structured training (mapathons) to strengthen national capacity and improve the consistency and reproducibility of land-use monitoring.

2–Sampling & Classification

Activity data were generated from a systematic national 4 × 4 km grid, covering Cameroon with 29,409 sample plots interpreted annually (2000–2023). Each 0.5 ha plot (71.7 × 71.7 m) was subdivided into 25 points (200 m^2 each), supporting consistent estimation of dominant class, forest threshold, and land-use transitions.

Interpretation in CEO used multi-source imagery—optical (Landsat 7–8, Sentinel-2, Planet) and radar (Sentinel-1)—supported by NDVI/NDFI indices to detect vegetation density and degradation, with Aqua MODIS (fire detection) and Google Earth Pro (validation) as auxiliary data.



The classification follows a two-tier system, based on the six IPCC land-use categories, disaggregated into classes adapted to national context.

IPCC Category	National Subcategories	Code
Forest	Dense humid forest	FDH
	Flooded and floodable forest	FIN
	Dry forest	FSE
	Planted forest	FPL
	Mangrove	FMA
Cropland	Annual crops	CAN
	Perennial crops	CPE
Grassland/Savanna	Grassland	PR
	Herbaceous savanna	SH
	Shrubby savanna	SA
Wetlands	Wetlands	H
Settlements	Settlements	E
Other land	Other land	A

Key considerations: the systematic 4 × 4 km grid provides robust national coverage but is less efficient for rare or high-impact classes. While large sample sizes (>29,000 samples) improve statistical precision, they may also propagate systematic interpretation bias at scale, especially when consensus processes introduce shared bias.

3–Decision Rules

A hierarchical decision tree reduced subjectivity in mixed plots: forest classification followed national definitions (≥30% continuous canopy cover within a 0.5 ha unit), while non-forest classes were assigned based on dominant cover and priority rules.

Decision Gate 1: Is continuous forest cover ≥30% (≥7.5 points)? — Yes → Classify as Forest; No ↓

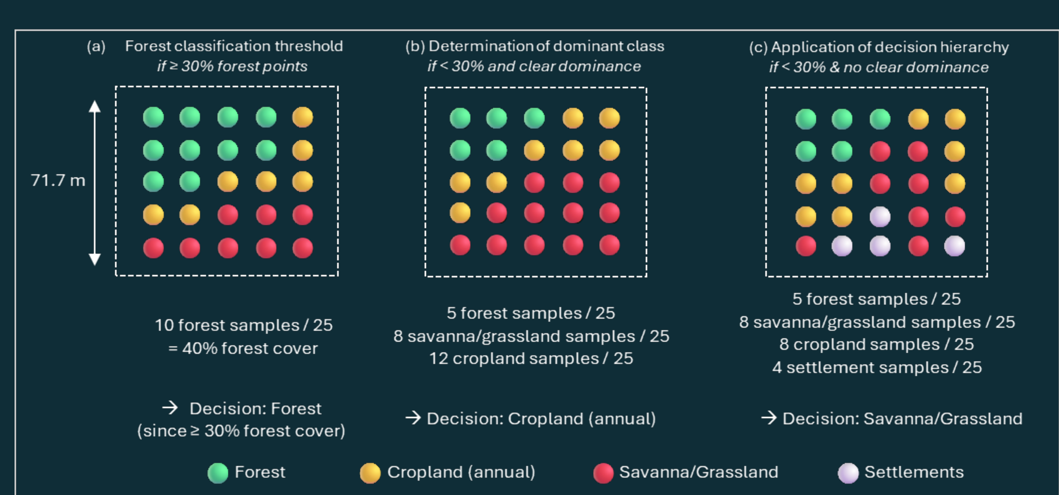
Decision Gate 2: Is there a clear dominant class among remaining points? — Yes → Assign to Dominant Class; No ↓

Apply Strict Priority Hierarchy
Forest > Grassland/Savanna > Cropland > Wetland > Settlement > Other

Filter Checkpoint: Plausibility Matrix
Excludes biophysically unlikely transitions, e.g. Mangrove → Dry Forest in one years

Interpretation was supported by technical sheets combining visual criteria, spectral indicators, and examples, with NDVI/NDFI aiding detection of subtle degradation.

Non-plausible transition rules flagged biophysically unlikely changes, improving temporal consistency.



Key considerations: class boundaries—especially degradation, savanna–forest mosaics, agroforestry, smallholder agriculture—can be interpretation-sensitive.

4–Integrated Framework for Activity Data, Emission Factors, QA/QC, & Capacity Building

Activity Data – A) A structured "mapathons" approach pre-production combined iterative capacity building with governance and bias tracking by bi-weekly reviews with ONACC and MINEPDED, with continuous updates to a bias registry & methodological guide:

1. Training & Alignment (Plots 1–150) → 2. Paired calibration (Plots 151–300) → 3. Independent Analysis (Plots 301–600) → 4. Operational Production (Plots 601+)

B) Production phase with QA/QC implementation for consistency between interpreters:

1. **Blind Double Interpretation (Plots 1–600)** Independent analysis by two experts, consensus if divergence.
2. **Control Re-sampling** ~30% of samples redistributed to verify cross-interpreter consistency.
3. **Automated Validation** System-level blocking of 'non-plausible transitions' (e.g., mangrove → dry forest).
4. **External Validation** Cross-referencing with national cartography and Global Forest Watch.

Emission Factors – GHG estimates follow IPCC guidance using a hierarchical approach: Tier 2 national data (e.g., Dees et al., 2018; ONACC, 2022; Gueguim et al., 2018), complemented by regional studies (e.g., NERF Gabon, 2021; Poulsen et al., 2020; Hubau et al., 2020; Kauffman & Bhomia, 2017) and Tier 1 IPCC defaults (2006; 2013 Wetlands Supplement; 2019 Refinement).

Key considerations: Uncertainty remains high due to limited national data (notably for biomass, soil carbon, degradation, and wood extraction); moving beyond Tier 1 through stronger inventories and data is key to reach Tier 2/3 and better reflect Cameroon's variability

5–Uncertainty Analysis

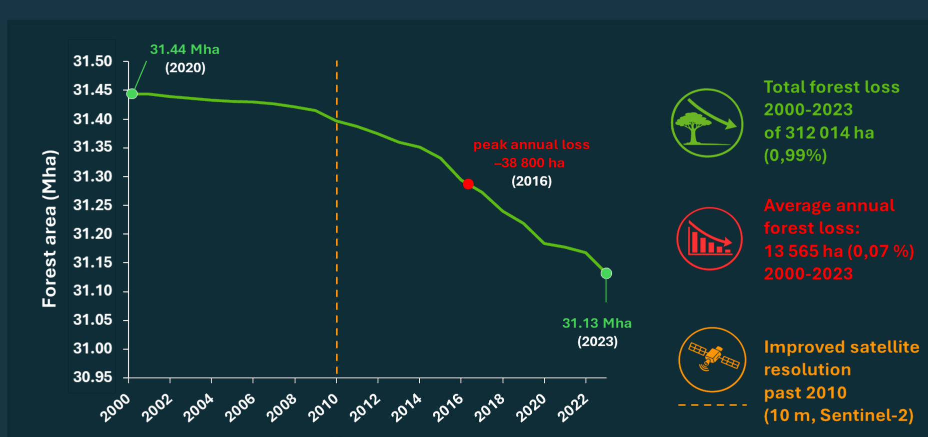
Approach:
IPCC Approach 2 with Monte Carlo simulations run in Argo (10,000 iterations; fixed seed for reproducibility).

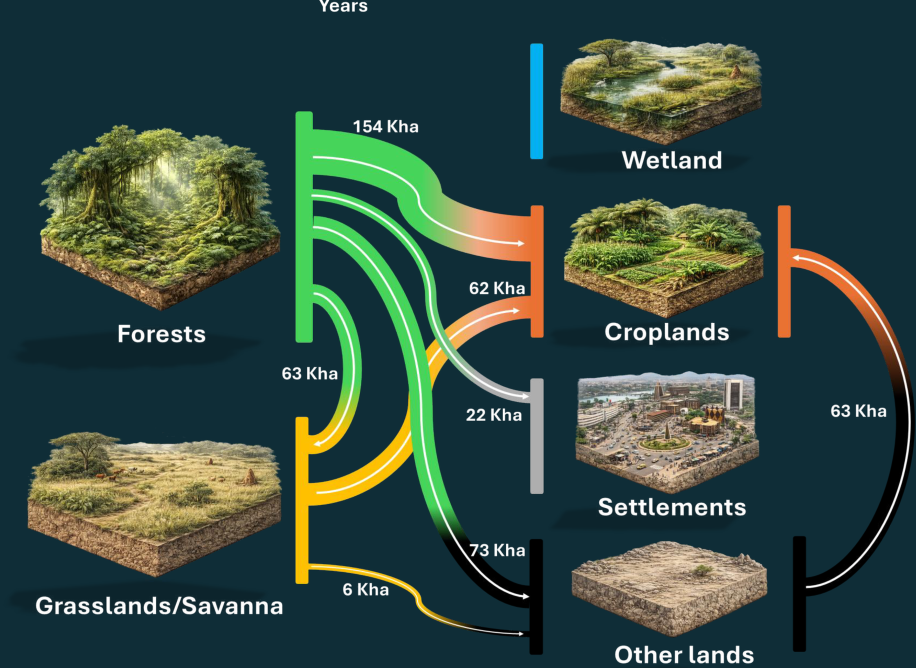
Error categorization:
- Random: sampling variability (e.g., forest = 0.8%)
- Systematic: interpretation bias compared with global & national maps (94.41% concordance; deforestation bias = 1.26%)

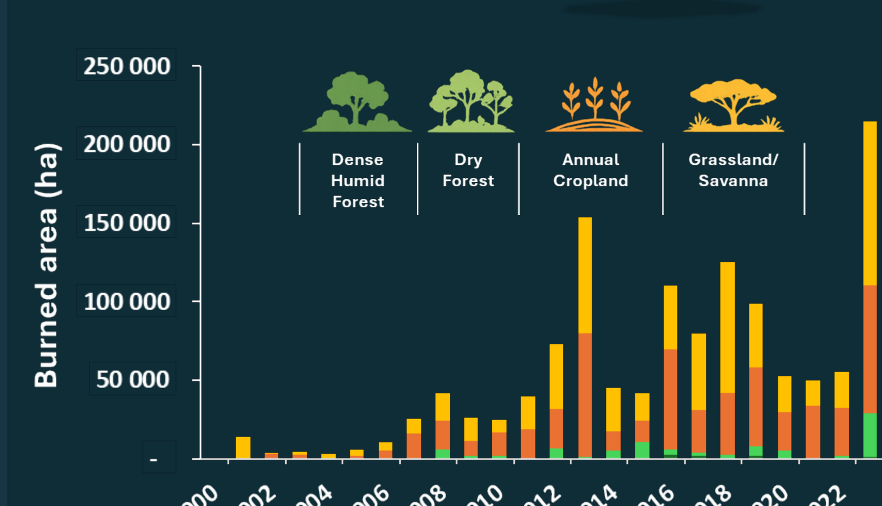
Monte Carlo Modeling

- Distributions: Inputs (activity data, emission factors, wood volumes) assigned triangular probability distributions (min., mode, max.)
- High uncertainty handling: Variables with >100% uncertainty used asymmetric distributions to avoid negative surface values
- Iterations: 10,000 runs per variable to ensure convergence to a stable distribution
- Reproducibility: Fixed random seed ensures external auditors can replicate results

Final Uncertainty Calculation
Uncertainty derived from simulation mean ± σ; 95% confidence interval (×1.96); example result: Net carbon absorption uncertainty (2020) = 32.1%.





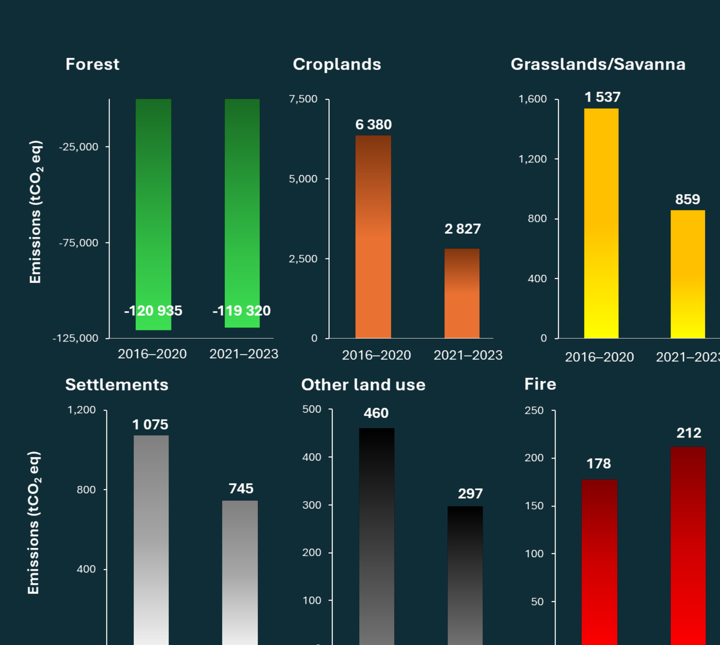


6–Key Results: Land use Dynamics and Emissions Estimates (2000–2023)

Forests dominate the landscape (67%) but declined by 312 Kha (–0.99%), mainly due to cropland expansion (154 Kha).

Fuelwood demand (>10 million m^3/year; FAOSTAT) and fire-driven agricultural expansion are the main drivers of degradation, alongside rising timber extraction (~3.4 to ~5.4 million m^3/year over 20 years; FAOSTAT, ITTO) and localized disturbances (flooding, small-scale mining).

Despite this, AFOLU remains a net carbon sink (–110,045 kt CO_2e in 2020; –105,203 kt CO_2e in 2023), though declining; emissions are driven by cropland and savanna, with CO_2 dominating (biomass) and CH_4 and N_2O mainly from fire.



7–Improvement Pathways

Priority actions to strengthen robustness and scalability:

Institutionalization & sustainability: The workflow remains project-based and must be sustained.
→ Embed within MINEPDED/ONACC with permanent teams, continuous training (mapathons), updated protocols, and a long-term improvement plan.

Classification sensitivity: Degradation, agroforestry, and savanna–forest mosaics remain difficult to distinguish.
→ Refine class definitions, expand technical sheets, and integrate confidence scoring.

Sampling & interpretation bias: The 4 × 4 km grid ensures robust coverage but is less efficient for rare classes; large samples (>29,000) may propagate shared bias. → Adopt stratified, spatially balanced sampling with class-/transition-specific precision targets; strengthen bias control (blind re-interpretation, confusion matrices, explicit disagreement in uncertainty).

Temporal consistency: Pre-2010 imagery limits detection of small-scale changes. → Prioritize Sentinel-era estimates and document confidence by period.

Emission factors: Reliance on Tier 1 increases uncertainty. → Develop Tier 2/3 data (NFI, soils, biomass, logging).

Uncertainty reporting: Sources remain aggregated. → Disaggregate sampling, interpretation, and EF uncertainties and target dominant contributors.