GIScience & Remote Sensing

ISSN: 1548-1603 (Print) 1943-7226 (Online) Journal homepage: www.tandfonline.com/journals/tgrs20

The importance of the sampling design in mapping woody cover in arid ecosystems

Felana Nantenaina Ramalason, Olivia Lovanirina Rakotondrasoa, Arthur Vander Linden, Guillaume Renard, Josoa R. Randriamalala, Nicolas J. Vereecken, Aina Razakamiaramanana , Mbolatiana F. Ranaivoharivelo , Sismondi Raholinarivo , Harifidy Rakoto Ratsimba , Jan Bogaert & Jean-François Bastin

To cite this article: Felana Nantenaina Ramalason, Olivia Lovanirina Rakotondrasoa, Arthur Vander Linden, Guillaume Renard, Josoa R. Randriamalala, Nicolas J. Vereecken, Aina Razakamiaramanana , Mbolatiana F. Ranaivoharivelo , Sismondi Raholinarivo , Harifidy Rakoto Ratsimba , Jan Bogaert & Jean-François Bastin (2026) The importance of the sampling design in mapping woody cover in arid ecosystems, GIScience & Remote Sensing, 63:1, 2658305, DOI: 10.1080/15481603.2026.2658305

To link to this article: https://doi.org/10.1080/15481603.2026.2658305

GISCIENCE & REMOTE SENSING
2026, VOL. 63, NO. 1, 2658305
https://doi.org/10.1080/15481603.2026.2658305

Taylor & Francis
by informa

RESEARCH ARTICLE

OPEN ACCESS Check for updates

The importance of the sampling design in mapping woody cover in arid ecosystems

Felana Nantenaina Ramalason[a,b], Olivia Lovanirina Rakotondrasoa[b,c], Arthur Vander Linden[a], Guillaume Renard[a], Josoa R. Randriamalala[b,c], Nicolas J. Vereecken[d], Aina Razakamiaramanana[b,c], Mbolatiana F. Ranaivoharivelo[b,c], Sismondi Raholinarivo[b,c], Harifidy Rakoto Ratsimba[b], Jan Bogaert[a] and Jean-François Bastin[a]

[a]University of Liège –TERRA Teaching and Research Centre, Gembloux Agro-Bio Tech. Passage des Déportés, Gembloux, Belgium; [b]Université d'Antananarivo - Ecole Supérieure des Sciences Agronomiques (ESSA-forêts), Mention Foresterie et Environnement, Ankatso, Antananarivo, Madagascar; [c]Ecole Doctorale Gestion des Ressources Naturelles et Développement (ED GRND), Université d'Antananarivo, Ankatso, Antananarivo, Madagascar; [d]Université libre de Bruxelles (ULB)- Agroecology Lab, Brussels, Belgium

ABSTRACT

Accurate calibration data remains a major constraint in ecological remote sensing, particularly in arid ecosystems where sparse and heterogeneous woody vegetation is challenging to detect. Despite advances in sensor technology and modelling algorithms, optimal calibration sampling design has received limited attention, with current approaches varying unsystematically (35–1,000 plots/1,000 km^2) without empirical benchmarks. Here, we develop and evaluate a quantitative framework for optimizing sampling strategies in remote-sensing models of woody cover through systematic comparison of (i) calibration data source (field surveys versus photointerpretation), (ii) spatial configuration (clustered versus dispersed), and (iii) sampling density. This integrated approach enables isolating the relative contributions of each design component to mapping accuracy—a critical gap in current remote sensing methodology.

We applied this framework to Sentinel-1, Sentinel-2, combined Sentinel-1 + 2, and AlphaEarth Foundations across Madagascar's arid southwest, validating predictions through spatial cross-validation using independent field plots. Photointerpretation substantially outperformed field-based calibration under typical arid-zone constraints ($R^2 = 0.88$, RMSE = 0.11 versus $R^2 = 0.46–0.66$, RMSE = 0.17–0.21). Performance saturated at 20.7–41.4 dispersed calibration plots per 1,000 km^2 across all predictors, beyond which gains became marginal. Dispersed strategies required half as many samples as clustered designs to achieve comparable accuracy, demonstrating that spatial distribution outweighs sample size. Once adequate sampling density and distribution were achieved, Sentinel-1 + 2 and AlphaEarth performed similarly ($R^2 \approx 0.85–0.86$), indicating that sampling design outweighs predictor complexity.

Our framework provides empirically derived operational thresholds (minimum: 15.5 plots/1,000 km^2; optimal: 20.7–41.4 plots/1,000 km^2) and a transferable methodology for determining optimal calibration densities in heterogeneous ecosystems where logistical constraints limit field sampling.

ARTICLE HISTORY
Received 13 January 2026
Accepted 7 April 2026

KEYWORDS
Sampling design; photointerpretation; remote sensing; arid ecosystems; woody vegetation

1. Introduction

Arid and semi-arid ecosystems cover 40.6% of Earth's land surface and support 2.3 billion people (UNCCD 2024), providing critical ecosystem services including wood production, carbon sequestration, and biodiversity conservation (D'Odorico et al. 2013; Feldman et al. 2024). Woody vegetation structures these ecosystems by influencing hydrological cycles, soil stability, and wildlife habitats (Eldridge et al. 2011). However, accelerated degradation driven by anthropogenic pressures—defined here as human-induced disturbances including land clearing, overgrazing, resource extraction and agricultural expansion (Reynolds et al. 2007)— and climate change makes monitoring essential for conservation and sustainable management (Feng and Fu 2013).

CONTACT Felana Nantenaina Ramalason rfelananantenaina@gmail.com; Jean-François Bastin jfbastin@uliege.be

Supplemental data for this article can be accessed online at https://doi.org/10.1080/15481603.2026.2658305.

Traditional field inventories provide accurate data but present major limitations: prohibitive costs, logistical constraints, and limited spatial coverage restrict large-scale application (McRoberts et al. 2018). Satellite remote sensing enables broad-scale monitoring but encounters specific challenges in arid environments. Sparse and heterogeneous vegetation, low biomass, short phenological cycles, and spectral confusion between senescent vegetation and bare soil complicate detection (Brandt and Hiernaux, 2016; Abel et al. 2019). Adaptive forms including deciduousness, leaf reduction, and waxy coatings further disrupt spectral detection. Low canopy density creates additional interpretation difficulties, as tree spectral signatures easily confound with exposed soil or shadows (Bastin et al. 2017; Brandt et al. 2020). Landscape fragmentation and deforestation further accentuate these challenges (De Marzo et al. 2021).

Recent advances in very high-resolution mapping have substantially improved our capacity to characterise woody vegetation at fine spatial scales. Global tree density estimation has been achieved at sub-kilometre resolution (Crowther et al. 2015), while sub-pixel decomposition methods including pixel-swapping and localised partial unmixing enable tree counting at landscape scale from medium-resolution imagery (Pouliot et al. 2018). Drone-based canopy mapping further extends these capabilities to centimetre-level resolution in semi-arid shrublands (Pérez-Luque et al. 2022). Deep learning applied to submetre imagery has enabled wall-to-wall mapping of over 1.8 billion individual tree crowns across dryland Africa (Brandt et al. 2020). Integrating Sentinel-1 (radar) and Sentinel-2 (optical) data offers improved detection through complementarity: radar ensures canopy penetration and moisture sensitivity, while optical data provide detailed spectral information (Dostálová et al. 2014; Acharki et al. 2021). Multi-temporal remote sensing analyses have further demonstrated the value of combining optical and spatial data for monitoring forest cover changes at regional scale (Valjarević et al. 2018). However, effectiveness depends critically on calibration data quality and representativeness (Ploton et al. 2020).

A critical methodological gap persists: despite widespread recognition that calibration sampling design influences model accuracy, current approaches vary unsystematically. Published studies report sampling densities ranging from 35 to 1,000 plots per 1,000 km^2 and plot sizes from 0.5 to 1 hectare, reflecting no consensus on optimal design (Bastin et al. 2017; Goldblatt, Rivera Ballesteros, and Burney 2017; Van Passel, De Keersmaecker, and Somers 2020). This uncertainty generates inefficiencies: some projects oversample at the expense of spatial coverage, while others undersample, compromising accuracy. Consequently, global woody cover maps present systematic biases in arid ecosystems (Ramalason et al. 2025b). Critically, no study has systematically quantified the relative contributions of calibration data source, spatial configuration, and sampling density to mapping accuracy across multiple sensor types—preventing development of evidence-based sampling guidelines.

Photointerpretation of very high-resolution imagery emerges as a promising intermediate solution between costly field campaigns and potentially biased global products (Pengra et al. 2007). This approach exploits human expertise to visually interpret complex vegetation patterns, enabling production of more numerous and representative samples while maintaining acceptable accuracy. Field plots provide high measurement precision but limited spatial coverage under typical logistical constraints in arid regions, while photointerpretation enables production of more numerous and spatially representative samples while maintaining acceptable accuracy. Concurrently, new pre-processed satellite datasets offer opportunities to improve ecological mapping. AlphaEarth Foundations, for example, combines spectral, textural, and topographic data through deep learning in standardised, directly exploitable formats (Brown et al., 2025).

Here, we address this methodological gap by developing an integrated quantitative framework for optimising calibration sampling design. Our approach advances beyond previous studies in three ways. First, we systematically isolate and quantify the independent effects of three calibration design components—data source (field versus photointerpretation), spatial configuration (clustered versus dispersed), and sampling density (10–505 plots)—through factorial experimental design with 100 iterations per configuration. Second, we evaluate these effects across four satellite predictor sets with varying complexity (Sentinel-1, Sentinel-2, Sentinel-1 + 2, AlphaEarth Foundations) to determine whether optimal sampling strategies depend on predictor sophistication. Third, we implement rigorous spatial cross-validation using independent field plots, ensuring performance estimates reflect true predictive capacity rather than overfitting to spatially autocorrelated data (Valavi et al. 2018; Ploton et al. 2020).

We apply this framework to Madagascar's arid southwest, encompassing dry forests and xerophytic thickets with high variability in vegetation structure and intense anthropogenic pressures, including deforestation linked to charcoal production, fuelwood extraction, and construction timber harvesting, as well as slash-and-burn agriculture (*tavy*) (Hänke et al. 2017; Randriamalala et al. 2022). This environmental heterogeneity provides a suitable testbed for evaluating sampling strategies across broad ecological gradients characteristic of global arid ecosystems.

Our objectives are to: (1) quantify how sampling design influences model accuracy relative to predictor complexity through systematic comparison; (2) identify minimal and optimal densities of calibration samples using multi-criteria convergence analysis; and (3) provide empirically derived operational thresholds and a transferable framework for designing calibration datasets under logistical and financial constraints. We test three hypotheses: (1) photo-interpreted samples produce more accurate maps than field inventories through better spatial representativeness; (2) performance increases with sample size up to an identifiable saturation threshold beyond which marginal gains diminish; and (3) multi-sensor integration improves accuracy progressively from single-sensor approaches to multi-dimensional composites, but only when calibration design is adequate.

This framework provides the first empirically derived quantitative guidelines for calibration sampling in arid ecosystem remote sensing, directly addressing the current lack of standardised approaches. The resulting operational thresholds and comparative methodology are transferable to other heterogeneous ecosystems where logistical constraints limit field sampling, offering practical guidance for balancing mapping accuracy against resource investment.

2. Materials and methods

2.1. Study area

We applied our framework to the Atsimo Andrefana region of southwestern Madagascar (4,826 km^2; 23°00'-24°30'S, 43°30'-44°30'E). The region experiences an arid to semi-arid climate (mean aridity index AI = 0.21; range: 0.13–0.36, following the UNCCD classification; (Trabucco and Zomer 2019; UNCCD 2024; Valjarević et al. 2025) with mean annual precipitation of 466 mm (range: 274–770 mm) and a mean annual temperature of 24.8 °C (2000-2022: CHIRPS Funk et al. 2015; ERA5-Land, Hersbach et al. 2020). The southern portions of the region are classified as arid (AI < 0.20) while the majority of the zone is semi-arid (AI 0.20–0.50). Rainfall is strongly seasonal, concentrated between December and March (January mean: 151 mm; August mean: 4 mm), with seven months receiving less than 30 mm per month. The hydrographic network is dominated by seasonal rivers with intermittent flow (MERIT Hydro; Yamazaki et al. 2019)). Topographic relief is moderate, with elevations ranging from near sea level to 598 m (mean: 162 m; SD: 129 m; Copernicus DEM GLO-30).

The region is increasingly affected by intensifying agricultural droughts and heat waves, threatening food security and endemic biodiversity. Although geographically less exposed to tropical cyclones than Madagascar's eastern coast, the southwest remains vulnerable to their indirect impacts—strong winds and intense rainfall episodes—in a context of high social vulnerability. Climate projections indicate a long-term trend toward reduced and more variable annual rainfall, increasing physiological stress on woody vegetation (IPCC, 2023). These climatic pressures underscore the urgency of robust and cost-effective woody cover monitoring frameworks for this region.

Vegetation comprises two co-existing structurally heterogeneous formations along an east-west gradient: (1) coastal xerophytic thickets, characterised by open structure and dominance of succulent species from endemic families Didiereaceae (*Didiera madagascariensis, Alluaudia procera)* and Euphorbiaceae (*Euphorbia stenoclada*); and (2) inland spiny forest, denser and more semi-deciduous, with a greater contribution of Fabaceae. This environmental heterogeneity provides a suitable testbed for evaluating sampling strategies across broad gradients.

2.2. Satellite data acquisition and preprocessing

2.2.1. Multi-sensor configurations

We used four satellite configurations to enable comparative assessment of predictors with variable complexity.

Sentinel-1 (S1): C-band SAR data were obtained from the Google Earth Engine (GEE) collection COPERNICUS/S1_GRD (Ground Range Detected, Level-1, Interferometric Wide Swath mode, VV and VH polarisations, 10 m resolution), offering partial canopy penetration and sensitivity to vegetation structure and moisture (Torres et al. 2012). A total of 10 scenes from both ascending and descending orbits were acquired over the study area (S1.2 Table S1).

Sentinel-2 (S2): Level 2 A surface reflectance data were obtained from the GEE collection COPERNICUS/S2_SR, atmospherically corrected by the ESA Sen2Cor processor. Bands B2 (blue, 490 nm), B3 (green, 560 nm), B4 (red, 665 nm), B8 (NIR, 842 nm), B11 (SWIR, 1,610 nm), and B12 (SWIR, 2,190 nm) were included, resampled to 10 m resolution. NDVI [(B8–B4)/(B8 + B4)] complemented spectral bands (Drusch et al. 2012). A total of 25 scenes were selected based on a cloud cover threshold of < 20% (S1.2 Table S2).

Sentinel-1 + 2 (S1 + S2): Combined configuration integrating radar structural sensitivity with optical spectral discrimination (Acharki et al. 2021).

AlphaEarth Foundations: Geospatial embedding model generating 64-band embeddings representing annual space-time trajectories (Brown et al., 2025). Each pixel embedding integrates spectral indices (NDVI, EVI, NDWI), textural metrics from optical and radar imagery(Sentinel-2 L1C, Landsat 8/9, and Sentinel-1 GRD imagery), topographic variables (elevation, slope, aspect, curvature from Copernicus DEM GLO-30), and spatialized climatic indices (derived from ERA5-Land), capturing ecosystem multidimensional complexity in analysis-ready format.

2.2.2. Temporal acquisition and preprocessing

All imagery was acquired during April 2022, corresponding to field data collection (S1.1 Sentinel-1 GRD data (VV and VH polarisations, Interferometric Wide Swath mode) were obtained from the COPERNICUS/S1_GRD GEE collection and composited as a monthly median for April 2022 (S1.1, S1.3). Thermal noise removal, radiometric calibration, and terrain correction are included in the standard GRD pipeline. Sentinel-2 surface reflectance imagery was obtained from COPERNICUS/S2_SR, filtered for cloud cover < 20% (CLOUDY_PIXEL_PERCENTAGE), and composited using a pixel-wise median over April 2022 acquisitions (S1.1, S1.3). AlphaEarth annual embeddings (64 bands, 2022) were obtained from GOOGLE/SATELLITE_EMBEDDING/V1/ANNUAL and mosaicked across the study area. All datasets were clipped to the study area and processed at 10 m resolution. Spectral and radar values were extracted at plot centroids.

2.3. Calibration data sources

2.3.1. Field data

Field sampling (April–May 2022) used stratified random design. We established 41 plots (30 × 30 m) distributed proportionally across strata defined by combining five woody cover classes (20% intervals) with four vegetation formations. Each site was selected to ensure structural homogeneity within a 200 m radius. Plots were geolocated using a Garmin GPSMAP 65 s (positional accuracy ± 1.5 m) and delineated using a Vertex hypsometer, with plots systematically located in homogeneous vegetation patches to minimise pixel-boundary effects.

Woody cover was measured using the Point-Intercept method adapted from the Australian Terrestrial Ecosystem Research Network protocol (White et al. 2012; Karl, McCord, and Hadley 2017). Within each plot, 54 points were systematically distributed along four orthogonal transects spaced 10 m apart, with points positioned every 2 m. Woody presence (≥1 m height) was recorded at each point by vertical projection. Woody cover was calculated as the proportion of positive points (S4.1, Figure S4.2).

2.3.2. Photo-interpreted data

We established 505 plots (50 × 50 m) through photo-interpretation of very high-resolution orthophotographs (2021–2022) using Collect Earth Online (Bey et al. 2016).

Sampling followed a stratified systematic unaligned design to capture the full woody cover gradient (S5.1, Figure S5.1). Strata were defined using a systematic grid with one random point per cell. Environmental variables including topography (elevation from NASADEM), climatic gradients (mean annual precipitation), and vegetation indices (NDVI, EVI, NDWI) were displayed as auxiliary contextual layers in

Collect Earth Online to support woody cover visual assessment during photo-interpretation (Colditz et al. 2011).

Each plot contained 49 observation points arranged in a 7 × 7 grid with approximately 7 m spacing (Figure 1, S5.2). Woody vegetation presence or absence at each point was determined through vertical photointerpretation using identical biophysical criteria as field data. This approach multiplied spatial sampling density by 12-fold while maintaining metric compatibility with field measurements.

Direct comparison between photo-interpreted and field-measured woody cover values for the 41 co-located field plots yields $R^2 = 0.775$, RMSE = 0.125, MAE = 0.100, and a mean relative error of 23.7% (regression slope = 0.92, bias = +0.042), confirming acceptable accuracy of photo-interpreted estimates relative to field measurements and supporting their use as calibration input under the logistical constraints typical of arid regions.

2.4. Comparative modelling and validation

Our analytical framework compared two calibration approaches across four satellite data sources using Random Forest algorithms (Figure 2). This design enabled systematic evaluation of how calibration data type, sample size, spatial distribution, and satellite data complexity influence woody cover mapping accuracy.

2.4.1. Random forest architecture

We developed eight Random Forest models independently: four calibrated exclusively with 41 field plots (one per satellite configuration) and four utilising all 505 photo-interpreted plots (one per configuration). In this supervised regression framework, each calibration plot provided a labelled observation—a field-measured or photo-interpreted wood cover value—paired with predictor values extracted as the mean of all pixels within each plot. For the Sentinel-1 + 2 configuration, predictors consisted of spectral and radar bands; for the AlphaEarth configuration, predictors were 64-dimensional embeddings derived from a deep learning model. Final models were then applied pixel-by-pixel across the entire study area to generate continuous wood cover prediction maps (0–100%). Random Forest was selected for its robustness against overfitting, ability to handle non-linear relationships, and resistance to correlated variables (Belgiu and Drăguţ 2016). The supervised learning process comprised three steps: (1) extraction of spectral or embedding predictor values per calibration polygon, calculated as the mean of all pixels within each plot footprint (30 × 30 m for field plots; 50 × 50 m for photo-interpreted plots); (2) supervised Random Forest model fitting on

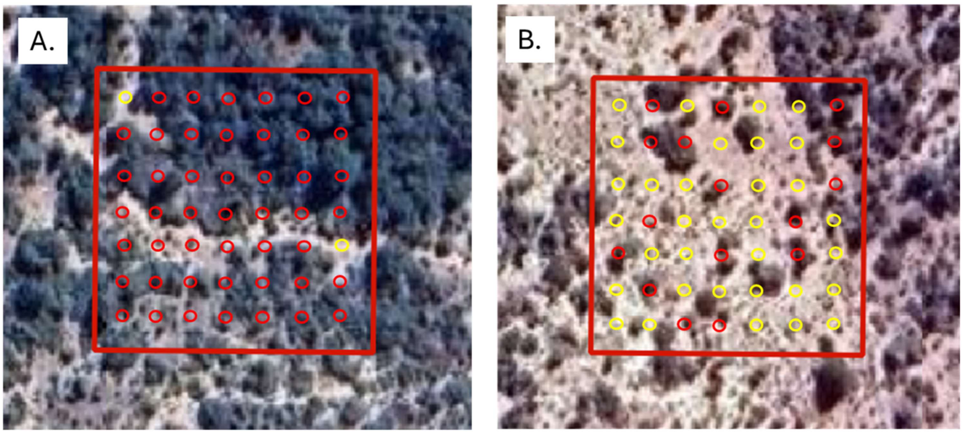

Figure 1. Photointerpretation of woody cover using very high-resolution orthophotographs. Examples of (A) dense woody cover (2/49 = 0.96) and (B) sparse woody cover (16/49 = 0.33). Each 50 × 50 m plot (red square) contains 49 observation points in a 7 × 7 grid. Red circles indicate woody vegetation presence; yellow circles indicate absence. A total of 505 plots were photo-interpreted from 2021–2022 orthophotographs using Collect Earth Online (Bey et al. 2016), following environmental stratification across topographic, climatic, vegetation, and edaphic gradients.

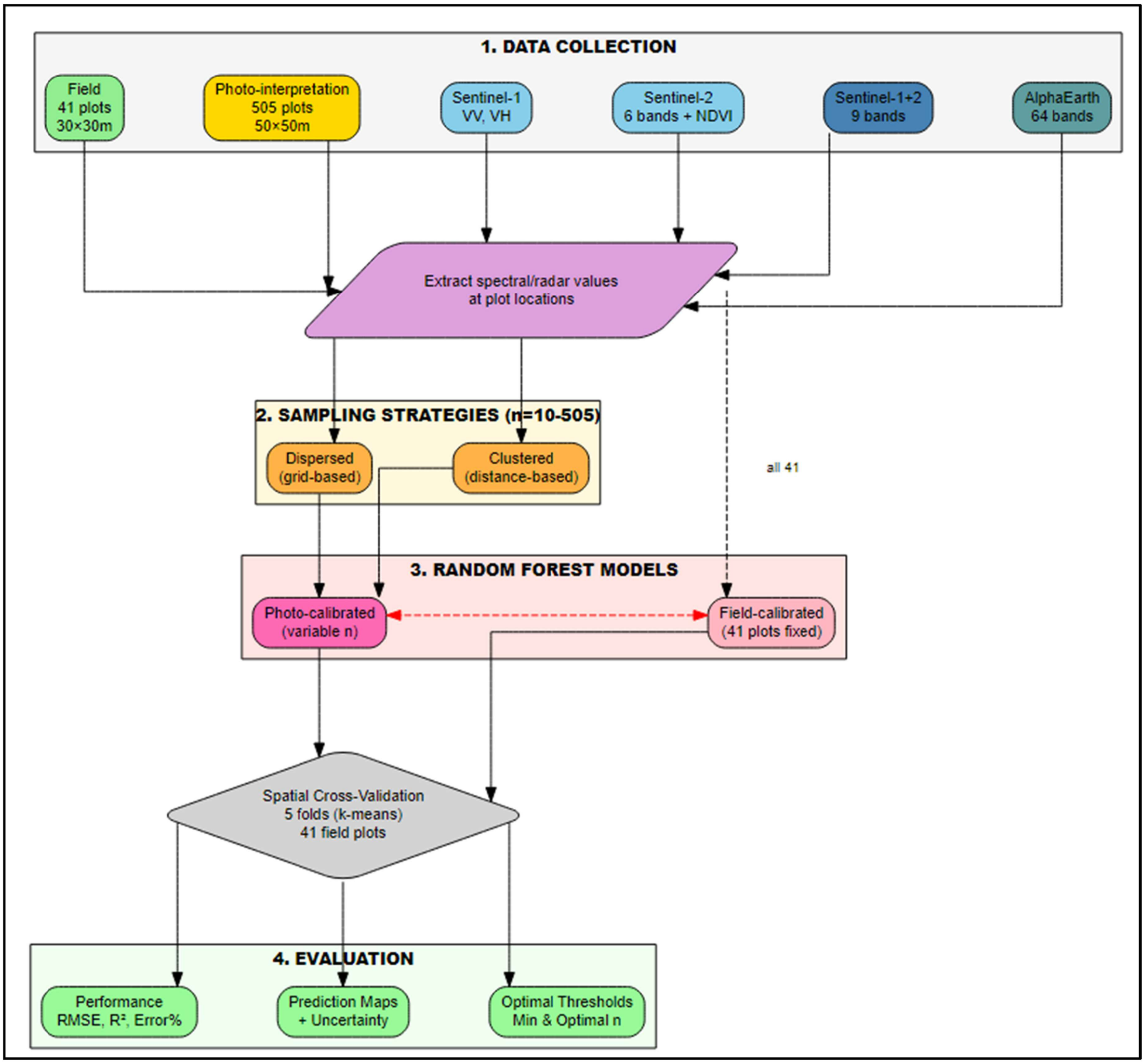

Figure 2. Methodological framework. The workflow integrates multi-source data collection (field plots, photointerpretation, and four satellite configurations: Sentinel-1(COPERNICUS/S1_GRD), Sentinel-2 (COPERNICUS/S2), Sentinel-1 + 2, and AlphaEarth), two sampling strategies (dispersed vs. clustered with $n = 10$-505), and two calibration sources (photo-calibrated with variable n and field-calibrated with 41 plots). Models are validated using spatial cross-validation where the 41 field plots are partitioned into 5 folds (k-means) and evaluated based on performance metrics (RMSE, R^2, Error%), prediction maps with uncertainty, and optimal thresholds and sample sizes.

(predictor, observed woody cover) pairs, with hyperparameter optimisation via grid search (ntree = 500, mtry = $\sqrt{p}$ where p is the number of predictors, nodesize = 5; Probst and Wright, 2019); (3) pixel-by-pixel spatial prediction across the entire study area to generate continuous woody cover maps (0–100%), following recommendations from Probst and Wright, 2019 (S3.1, S3.2, S3.3, Table S3.1). The response variable underwent square-root transformation for variance stabilisation before modelling, with back-transformation for evaluation (S3.3, S6.1). Variable importance was assessed using Mean Decrease in MSE, a metric inherent to the Random Forest algorithm (Breiman 2001); detailed results are provided in Supplementary Figure S3.3.

2.4.2. Validation protocols

We implemented spatial cross-validation by partitioning the 41 field plots into five mutually exclusive geographic blocks through k-means clustering of coordinates (k = 5, nstart = 25), maximising geographic separation between folds (Valavi et al. 2018) (S2.1, S2.3, Figure S2.1). Spatial independence was verified using Moran's I test ($I = -0.025$, $p = 0.5$), confirming absence of residual spatial autocorrelation between folds (S2.2). These spatial folds, defined from field plots, were consistently applied to evaluate both field-based and photointerpretation-based models, ensuring methodological consistency for direct performance comparison.

2.5. Spatial and statistical analyses

Model performance was assessed through three complementary metrics. First, we quantified prediction accuracy using spatial cross-validation on the 41 field plots, calculating RMSE, relative error, and R^2 (S7, S6.1). Second, we examined distributions of predicted values across the study area. Third, we quantified spatial uncertainty through coefficient of variation across the five folds for each pixel (Zhang et al., 2005; Nuzzo 2019) (S7.4). Systematic comparative analysis between satellite configurations enabled identifying relative influence of predictive data complexity versus calibration quality and density.

2.6. Sensitivity analysis to sampling density

2.6.1. Experimental design

To determine optimal photo-interpreted sampling density thresholds, we tested thirteen sample sizes: $n \in \{10, 25, 50, 75, 100, 150, 200, 250, 300, 350, 400, 450, 505$ plots$\}$, designed to capture the transition from minimal to complete sampling, with densification around expected transition zones (Mellor et al. 2015).

For each sample size below 505, we performed 100 iterations with different random seeds to quantify sampling variability and estimate performance confidence intervals (Olofsson et al. 2014) (S6.2). This analysis was repeated for each satellite configuration.

2.6.2. Contrasting spatial strategies

We compared two spatial distribution strategies. **Dispersed sampling** overlaid a regular grid and randomly selected one plot per cell, maximising environmental representativeness. **Clustered sampling** randomly selected one central point then selected nearest plots based on Euclidean distance, simulating logistical constraints typical of conservation programmes (Wang, Gertner, and Anderson 2005). Each combination of sample size and spatial strategy was systematically tested with all satellite configurations.

2.7. Evaluation of metrics and critical threshold identification

2.7.1. Performance metrics

We used three complementary metrics: R^2 quantified variance explained, RMSE measured average prediction error magnitude, and mean relative error normalised errors relative to observed values, facilitating comparisons across cover ranges.

2.7.2. Optimal threshold identification

Sampling thresholds were identified based on model performance and stability, evaluated through repeated calibrations for each sample size. Following the uncertainty-aware framework of Olofsson et al. (2014), the analysis accounts for performance variability to ensure statistically comparable estimates across configurations.

The minimal threshold corresponds to the smallest sample size meeting acceptable criteria: $R^2 \geq 0.75$, RMSE efficiency $\geq 85\%$, and coefficients of variation below 15% for both metrics. The optimal threshold identifies the point where additional plots yield marginal gains, determined using stricter criteria: $R^2 \geq 0.84$, RMSE efficiency $\geq 92\%$, coefficients of variation below 12%, and average RMSE improvement below 2% across subsequent sample sizes. Thresholds were identified independently for each satellite configuration and spatial design. Sample sizes were normalised as sampling density (plots per 1000 km^2) relative to the study area (4,826 km^2).

3. Results

3.1. Spatial representativeness outweighs measurement precision

Direct comparison between calibration approaches—field measurement precision (41 plots) versus photo-interpretation spatial density (505 plots)—revealed substantial performance differences when validated against identical independent field data (Figure 3). Although each photo-interpreted observation relies on visual assessment rather than direct field measurement, their 12-fold greater spatial density (10.5 versus 0.85 plots/100 km^2) consistently produced superior models across all satellite configurations. Using Sentinel-1 + 2, photointerpretation achieved $R^2 = 0.88$ (RMSE = 0.11) compared to $R^2 = 0.66$ (RMSE = 0.17) for field calibration—a 35% error reduction (Figure 3A). This performance gap widened with AlphaEarth: Photointerpretation maintained $R^2 = 0.88$ (RMSE = 0.10) while field calibration declined to $R^2 = 0.46$ (RMSE = 0.21) (Figure 3C).

The relationship between observed and predicted values differed markedly between approaches. Photointerpretation produced near-linear relationships with regression slopes of 1.09 (Sentinel-1 + 2) and 0.96 (AlphaEarth), approximating ideal 1:1 correspondence (Figure 3A, C). Field calibration yielded slopes of only 0.51 (Sentinel-1 + 2) and 0.27 (AlphaEarth), indicating systematic underestimation of woody cover. This compression became particularly acute above 60% cover, where field-calibrated predictions plateaued regardless of observed values.

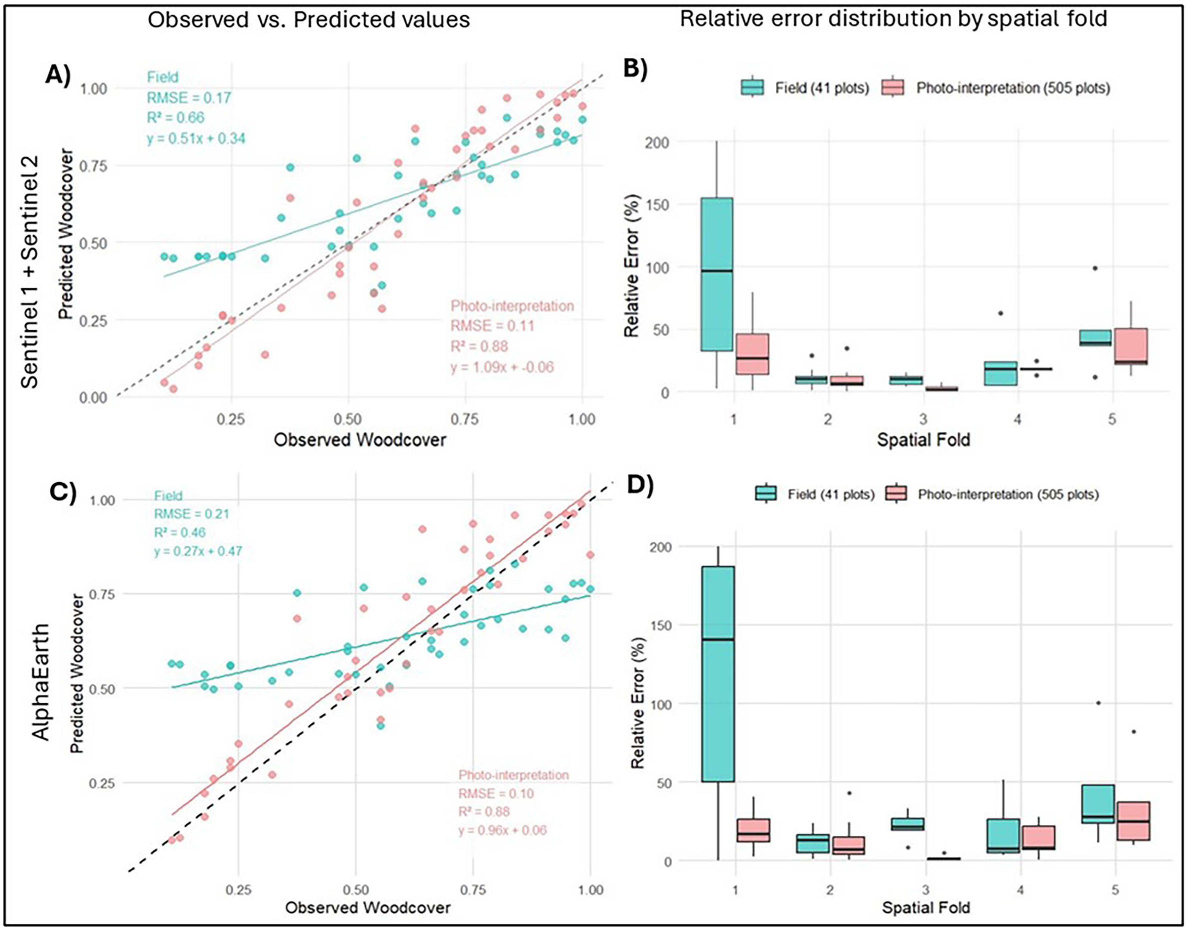

Figure 3. Performance comparison between field and photointerpretation calibration for mapping woody cover with Sentinel-1 + 2 (A, B) and AlphaEarth (C, D). (A, C) Observed vs. predicted scatter plots demonstrate photointerpretation superiority (505 plots: RMSE = 0.11, $R^2 = 0.88$ for Sentinel-1 + 2; RMSE = 0.10, $R^2 = 0.88$ for AlphaEarth) over field calibration alone (41 plots: RMSE = 0.17, $R^2 = 0.66$ for Sentinel-1 + 2; RMSE = 0.21, $R^2 = 0.46$ for AlphaEarth). (B, D) Relative error distribution by spatial validation blocks reveals reduced and stable variability for photointerpretation calibration versus field approaches for both satellite data sources.

Cross-validation revealed contrasting spatial stability patterns. Field calibration exhibited high variability among spatial folds, with median relative errors ranging from 20% to 150% across blocks (Figure 3B, D). This instability intensified with predictor complexity: AlphaEarth produced wider error distributions (10–190%) than Sentinel-1 + 2 (20–150%), suggesting insufficient sample size to constrain high-dimensional predictor space. Photointerpretation maintained stable median relative errors of 15–35% across all spatial blocks and satellite configurations, demonstrating robust spatial generalisation with 505 calibration plots.

3.2. Calibration density determines spatial pattern fidelity

Predicted woody cover maps differed qualitatively between calibration approaches (Figure 4). Field-calibrated maps (41 plots) displayed spatially homogeneous patterns with most predictions concentrated between 40–60% woody cover. Distribution histograms showed unimodal peaks centred at 50%, failing to reproduce the bimodal distribution observed in field data (primary mode: 20–30%; secondary mode: 60–70%) (Figure 4E, G). This compression toward intermediate values occurred with both Sentinel-1 + 2 (Figure 4A, E) and AlphaEarth (Figure 4C, G). Prediction uncertainty was high (30–40%) and spatially concentrated in coastal zones, expanding more broadly under AlphaEarth (Figure 4I, K).

Photo-interpretation-calibrated maps (505 plots) captured spatial heterogeneity through complex mosaics alternating between high-cover (70–90%) and low-cover (10–30%) zones (Figure 4B, D). Predicted distributions exhibited clear bimodality matching field observations, with primary peaks at 20–30% and secondary peaks at 60–70% (Figure 4F, H). This distributional fidelity persisted across both satellite configurations. Prediction uncertainty remained low (10–25%) and spatially uniform, with localised increases only in ecotonal transition zones (Figure 4J, L).

3.3. Sampling strategy and size control model convergence

Systematic variation of photo-interpretation sample size revealed distinct convergence trajectories by sampling strategy (Figure 5). Under dispersed sampling, performance metrics stabilised rapidly: R^2 exceeded 0.80 at 100–150 plots and plateaued at 0.88 beyond 200 plots for both Sentinel-1 + 2 and AlphaEarth. RMSE declined from 0.20 (10 plots) to 0.11–0.12 (>150 plots), while relative error decreased from 30% to 20% over the same range. Marginal gains beyond 200 plots were minimal (<2% R^2 improvement per 50-point increment).

Clustered sampling exhibited slower convergence with greater variability. Between 10 and 100 plots, R^2 fluctuated broadly (0.50–0.70) with large error bars indicating high sensitivity to specific plot locations. Performance approached dispersed sampling levels only at 400–500 plots ($R^2 = 0.85–0.88$, RMSE = 0.11–0.12), requiring 2–2.5 × more observations to achieve equivalent accuracy.

Satellite configuration comparisons showed differential data requirements. Sentinel-1 alone plateaued at $R^2 = 0.68$ regardless of sample size, confirming fundamental limitations of single-polarisation radar for woody cover estimation. Sentinel-2 required 25–30% more plots than Sentinel-1 + 2 to reach comparable performance, demonstrating synergistic value of optical-radar fusion. AlphaEarth and Sentinel-1 + 2 converged to similar asymptotic performance ($R^2 = 0.88$, RMSE = 0.11) from 150 plots onward under dispersed sampling, indicating limited marginal benefit of foundation model embeddings once adequate sampling is achieved.

3.4. Operational thresholds depend on spatial distribution

Quantitative threshold identification revealed consistent patterns across satellite configurations (Table 1). For dispersed sampling, minimum acceptable performance ($R^2 \geq 0.82$, RMSE ≤ 0.14) emerged at 15.5 plots/1000 km^2 (75 plots). Optimal performance balancing accuracy and sampling effort ($R^2 \geq 0.85$, RMSE ≤ 0.11) occurred at 20.7–41.4 plots/1000 km^2 (100–200 plots), beyond which cost-benefit ratios declined.

Clustered sampling required substantially higher densities. Minimum thresholds reached 31.1–41.4 plots/1000 km^2 (150–200 plots)—double that of dispersed sampling—while optimal thresholds

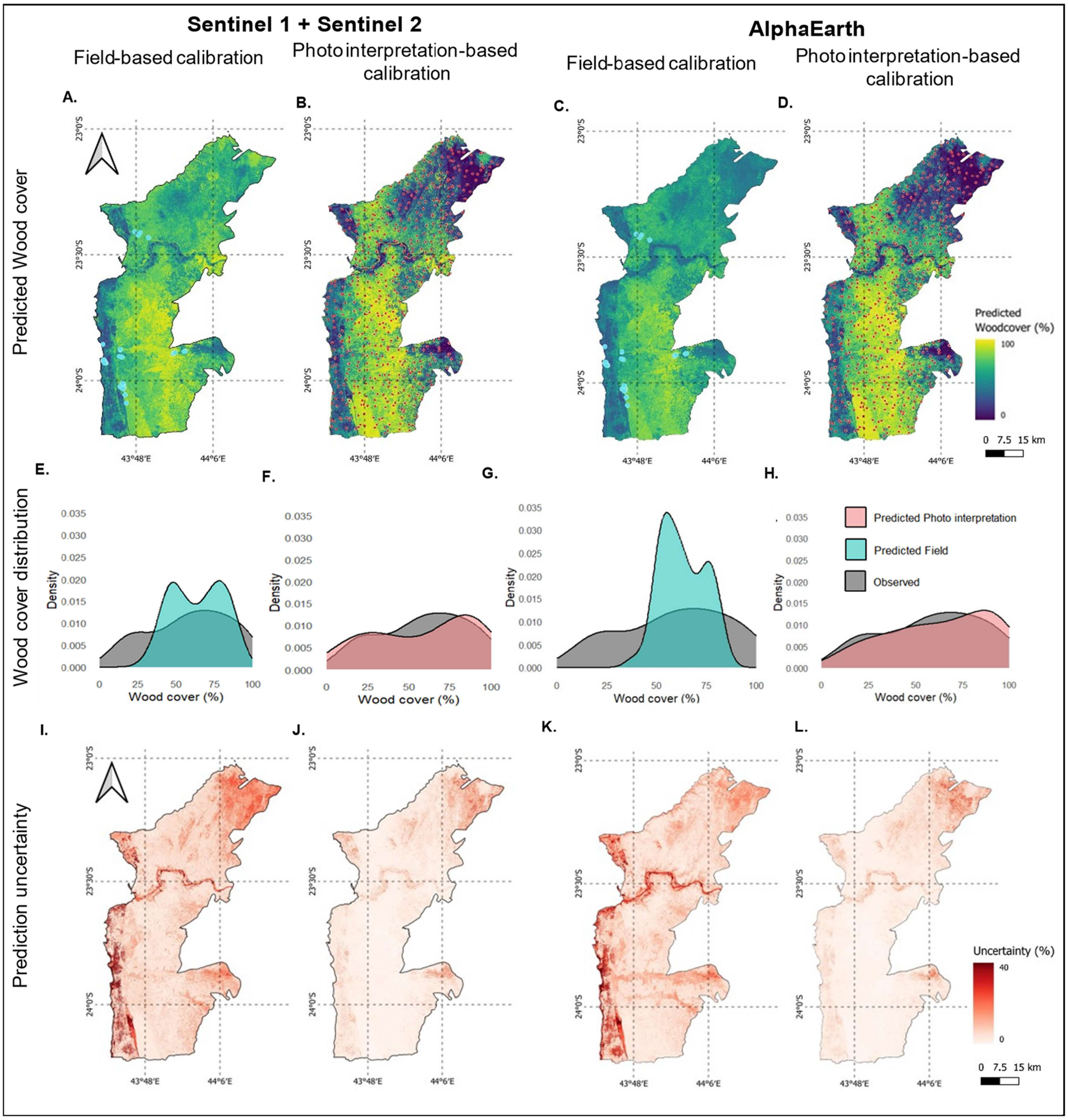

Figure 4. Spatial predictions of woody cover, frequency distributions, and prediction uncertainty. (A–D) Predicted woody cover maps, (E–H) frequency distributions comparing observed (grey) and predicted values (teal: field-based calibration; pink: photo interpretation-based calibration), and (I–L) prediction uncertainty maps for field-based calibration (A, C, E, G, I, K) and photo interpretation-based calibration (B, D, F, H, J, L) using Sentinel-1 + 2 (A, B, E, F, I, J) and AlphaEarth (C, D, G, H, K, L). Colour scales indicate woody cover percentage (0–100%) and prediction uncertainty (0–40%). North arrow shown in panels A and I, scale bar shown in panels D and L apply to all map panels. Geographic coordinate system: WGS 84 (EPSG:4326).

extended to 51.8–62.2 plots/1000 km^2 (250–300 plots). These elevated requirements reflect spatial redundancy inherent to clustered designs, where proximate observations provide diminishing marginal information.

Satellite configuration had minimal impact on threshold locations: Sentinel-1 + 2 and AlphaEarth differed by only 1–3% R^2 at equivalent sample sizes, with threshold densities varying within 10 plots/1000 km^2. Notably, minimal photo-interpretation sampling (75 dispersed plots; R^2 = 0.82) surpassed conventional field calibration (41 plots; R^2 = 0.46–0.66), demonstrating superior efficiency for resource-constrained applications.

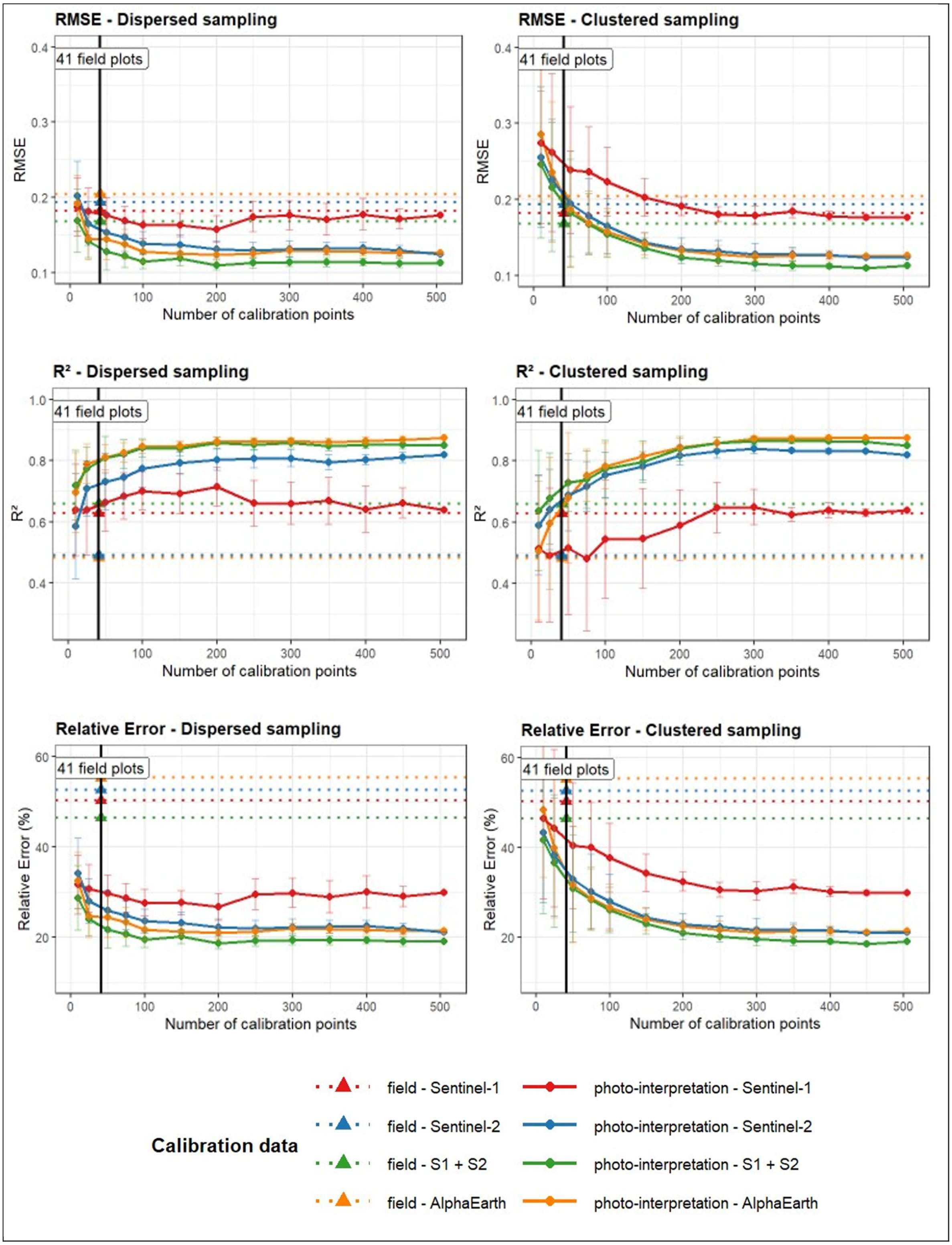

Figure 5. Performance metrics across photo interpretation sample sizes and sampling strategies. RMSE (top), R^2 (middle), and relative error (bottom) for dispersed (left) and clustered (right) sampling using Sentinel-1, Sentinel-2, Sentinel-1 + Sentinel-2, and AlphaEarth. Dashed lines with triangles: field-based calibration; solid lines with circles: photo interpretation-based calibration. Vertical line indicates baseline field calibration (41 plots). Error bars show standard deviations over 100 iterations.

Table 1. Minimum and optimal photointerpretation calibration point thresholds to obtain performant models according to RMSE, R^2, and relative error (%) metrics.

Distribution	Source	Threshold	Plots/1000 km^2	*N* plots	RMSE	R^2
Dispersed	Sentinel-1 + 2	Minimal	15.5	75	0, 12	0, 82
		Optimal	41.4	200	0, 11	0, 85
	AlphaEarth	Minimal	15.5	75	0, 14	0, 82
		Optimal	20.7	100	0, 14	0, 84
Clustered	Sentinel-1 + 2	Minimal	41.4	200	0, 12	0, 83
		Optimal	62.2	300	0, 13	0, 86
	AlphaEarth	Minimal	31.1	150	0, 14	0, 81
		Optimal	51.8	250	0, 13	0, 85

4. Discussion

4.1. Superiority of photointerpretation for model calibration

Models calibrated with 505 photo-interpreted plots achieved substantially higher performance (R^2 = 0.88, RMSE = 0.11) compared to those calibrated with 41 field plots (R^2 = 0.46–0.66, RMSE = 0.17–0.21) using spatially independent clusters via k-means spatial clustering for validation, representing a 35–90% improvement depending on satellite configuration. The superior performance of photointerpretation reflects a fundamental calibration principle: spatial representativeness outweighs measurement precision for machine learning models. Sparse field plots (41) created environmental gaps forcing model extrapolation (slopes = 0.27–0.51), while dense photointerpretation (505) enabled interpolation (slopes = 0.96–1.09). For context, Dynamic World—a globally trained deep-learning land cover product—achieved R^2 = 0.70 on the same 41 independent field plots (Brown et al. 2022; Ramalason et al. 2025b) markedly lower than our locally calibrated models (R^2 = 0.88), demonstrating the substantial benefit of regionally tailored calibration. Prior to this study, woody cover mapping in southwestern Madagascar had been limited to discrete land cover classification rather than continuous cover estimation (Ramalason, Rakotondrasoa, and Bastin 2025a), underscoring the novelty of the spatially continuous approach developed here.

Spatial representativeness translates into remarkable cross-validation stability: relative errors remain homogeneous (10–50%) regardless of geographic fold tested, while field calibration exhibits strong instability with errors ranging from 5% to 190% across folds, confirming results of Tsendbazar et al. (2021), who demonstrated that spatial representativeness outweighs raw sample quantity for large-scale mapping.

Direct validation of photo-interpreted against field-measured woody cover values for the 41 co-located plots confirms acceptable measurement accuracy (R^2 = 0.77, RMSE = 0.12, slope = 0.92, bias = +0.042), with the slight positive bias consistent with the known tendency of VHR imagery to capture canopy extent more completely than the point-intercept method. This demonstrates that photo-interpreted data provides both adequate spatial coverage and acceptable individual measurement quality for model calibration.

Our results align with the growing use of photointerpretation for ecological mapping. Provencher-Nolet, Bernier, and Lévesque (2014) achieved 85% accuracy for forest habitat classification using aerial photographs, while Pérez-Luque et al. (2022) demonstrated strong correlation (R^2 = 0.81) between drone-based estimates and field sampling in Mediterranean shrublands. Francis et al. (2023) showed that integrating drone and satellite data enables precise monitoring of vegetation changes at landscape scale. Advances in deep learning (Symeonakis et al. 2022) and increasing accessibility of cloud platforms (Guida-Johnson et al. 2021) further enhance operational potential.

These performances do not diminish the intrinsic value of field surveys, which provide high point-level accuracy and complementary ecological information (floristic composition, vertical structure). However, in vast arid landscapes where logistical and financial constraints limit field sampling density (McRoberts et al. 2018), photointerpretation constitutes a pragmatic alternative, provided rigorous validation and quality control protocols are implemented.

4.2. Critical thresholds and spatial distribution of samples

The identification of critical sampling density thresholds represents a major methodological contribution for designing calibration campaigns. Three key findings emerge.

First, spatial distribution dominates: clustered sampling requires twice as many points as dispersed sampling to achieve equivalent performance. For spatially dispersed sampling, we identified a minimum threshold of 15.5 plots/1000 km^2 (75–100 plots) ensuring acceptable accuracy ($R^2 = 0.82$, RMSE ≤ 0.14), and an optimal threshold of 20.7–41.4 plots/1000 km^2 (100–200 plots) maximising cost-benefit ratio ($R^2 = 0.85$–0.86, RMSE = 0.11–0.12). Beyond this threshold, performance gains become marginal (<2% per 50 additional plots), suggesting saturation of predictive capacity. Clustered sampling requires substantially higher densities: 31.1–41.4 points/1000 km^2 (150–200 plots) for minimum threshold and 51.8–62.2 points/1000 km^2 (250–300 plots) for optimal threshold—systematically double that of dispersed sampling.

Second, AlphaEarth provides negligible advantage over Sentinel-1 + 2 (1–3% R^2 gain) when adequate calibration samples are available, demonstrating that sample design outweighs satellite data sophistication.

Third, minimal photointerpretation (75 dispersed points; $R^2 = 0.82$) outperforms 41-point field calibration ($R^2 = 0.46$–0.66), confirming superior cost-effectiveness for resource-limited monitoring programmes. This difference persists regardless of satellite data source, demonstrating that spatial distribution constitutes the dominant factor affecting model performance. Even with minimal but dispersed photo-interpreted sampling, performance substantially exceeds that obtained with field plots, underscoring critical importance of geographic coverage.

These results complement the work of Mellor et al. (2015), who observed saturation of Random Forest performance beyond 200–300 samples for land cover mapping, and Qu et al. (2024), who identified performance plateaus for forest mapping from approximately 150 samples. Our contribution extends these observations by precisely quantifying thresholds for different spatial strategies and data configurations, providing directly transferable operational benchmarks.

Normalisation as surface density (plots/1000 km^2) facilitates extrapolation to other geographic contexts, although the thresholds identified here are specific to this biome and the Random Forest modelling algorithm. More fragmented or structurally complex ecosystems may require higher densities to capture environmental variability (Stehman and Foody 2019). Conversely, more spatially uniform ecosystems are expected to reach performance saturation at lower densities. Nevertheless, these thresholds provide a solid foundation for dimensioning sampling efforts in operational monitoring programmes, particularly for government agencies, NGOs, and local projects facing budget constraints.

4.3. Multi-sensor complementarity: gains conditional on sample adequacy

Comparative analysis of satellite data sources reveals nuanced results challenging our hypothesis of systematic progressive improvement with predictor complexity. Individual sensors showed distinct limitations: Sentinel-1 alone remained constrained regardless of sample size (R^2 max = 0.68), confirming that radar data, while valuable for capturing structural information including canopy height, density, and moisture, cannot fully represent spectral variability necessary for accurate woody cover estimation in heterogeneous arid landscapes. Sentinel-2 performed better but required 25–30% additional calibration plots compared to Sentinel-1 + 2 to achieve equivalent performance.

The Sentinel-1 + 2 combination demonstrated clear advantages with limited samples, achieving $R^2 = 0.66$ with only 41 field plots compared to lower values for individual sensors. This synergy—radar providing structural metrics and optical data capturing biochemical variability including chlorophyll and vegetation indices—proves particularly valuable in arid ecosystems where seasonal phenology and sparse vegetation complicate spectral interpretation, confirming findings by Frison et al. (2018) and Reiche et al. (2018) in tropical forests, and recent studies in arid zones (Karakizi et al. 2023).

However, AlphaEarth contradicted expected progressive improvement. With adequate photo-interpreted sampling (150–200 plots), both AlphaEarth and Sentinel-1 + 2 converged toward nearly identical performance ($R^2 = 0.88$, RMSE = 0.11), with only 1–2% marginal superiority for AlphaEarth. More strikingly, with limited field calibration (41 plots), AlphaEarth substantially underperformed ($R^2 = 0.46$ vs 0.66 for Sentinel-1 + 2).

This performance collapse illustrates the curse of dimensionality (Hawkins 2004). AlphaEarth integrates dozens of variables including spectral indices, textural metrics, and topographic derivatives, creating a high-dimensional predictor space. With only 41 observations, Random Forest cannot stably estimate relationships between these numerous predictors and woody cover, resulting in overfitting where the model memorises training data rather than learning generalisable patterns. Conversely, Sentinel-1 + 2 combines fewer fundamental variables including radar polarisations, spectral bands, and NDVI, requiring less calibration data for robust parameter estimation.

Variable importance analysis further revealed that VH backscatter dominated Sentinel-1 + 2 predictions under both calibration sources, while AlphaEarth showed a more diffuse importance distribution under photointerpretation, reflecting broader predictor exploitation with larger training sets (Figure S3.3).

These findings demonstrate that sample design—not satellite data sophistication—primarily determines mapping accuracy in resource-constrained contexts. Investing in spatially representative photo-interpreted samples combined with freely accessible dual-sensor data (Sentinel-1 + 2) proves more effective than limited field calibration paired with complex multi-dimensional composites, confirming observations by (Foody 2009; Mellor et al. 2015) that sample size and spatial distribution outweigh predictor complexity.

4.4. Limitations and methodological perspectives

Five principal limitations warrant explicit acknowledgement.

First, photointerpretation relies on visual judgement by human operators, introducing inter-observer variability that can affect result consistency. In this study, two operators followed a standardised protocol, but no formal inter-rater reliability assessment was conducted. Insufficient image resolution further amplifies this uncertainty in open and sparse woody formations where woody-herbaceous discrimination becomes ambiguous. Future studies should systematically cross-validate subsamples across multiple operators to quantify and minimise these biases.

Second, the limited number of field plots (41) constrains external validation of photo-interpreted models. Direct comparison between photo-interpreted and field-measured woody cover values for the 41 co-located plots ($R^2 = 0.775$, RMSE = 0.125, slope = 0.92, bias = + 0.042) partially addresses this limitation by quantifying the accuracy of photo-interpreted data as calibration input; a fully independent validation set of larger size would nonetheless further strengthen conclusions. However, we implemented rigorous spatial cross-validation via k-means spatial clustering, which accounts for spatial autocorrelation by ensuring that validation clusters are geographically distinct from training clusters. This approach is fundamental, as Ploton et al. (2020) revealed poor predictive performance for large-scale ecological models when spatial autocorrelation is ignored. Recent methodological developments, such as the CAST package (Meyer et al. 2024) and recommendations by Stock (2025), reinforce the necessity of these approaches for robust evaluation of model spatial generalisation.

Third, the calibration density thresholds identified here (15.5–41.4 plots per 1,000 km^2) are specific to the Malagasy spiny forest biome and to the Random Forest algorithm, whose predictive performance varies with ecosystem context (Belgiu and Drăguţ 2016). Structurally heterogeneous ecosystems will require higher calibration densities, whereas spatially homogeneous ecosystems are expected to reach performance saturation at lower densities (Fassnacht et al. 2014). Furthermore, alternative algorithms—notably Support Vector Machines or deep learning approaches potentially better suited to exploiting the high-dimensional embedding space of AlphaEarth—may yield different threshold values (Fassnacht et al. 2016). Validating this framework with alternative algorithms therefore constitutes a priority for future research.

4.5. Implications for sustainable management of arid ecosystems

These results have direct applications for sustainable management of woody resources. Fine-scale mapping of woody cover enables identification of declining areas (deforestation, overgrazing), remnant vegetation cores requiring priority conservation, and areas favourable for natural or assisted regeneration. (Brandt et al. 2020) revealed an unexpected number of trees in Sahel drylands using high-resolution imagery, substantially revising carbon stock estimates. Sengani, Ramoelo, and Archer (2023) highlighted the importance of woody cover mapping for early detection of degradation in semi-arid regions.

The proposed methodology integrates with Land Degradation Neutrality (LDN) objectives, providing spatially explicit information to guide restoration actions (Bastin et al. 2019), track international commitments, and justify investments. The accessibility of platforms like Collect Earth Online facilitates implementation of participatory protocols where local communities can contribute to data collection, strengthening local ownership and sustainability of monitoring systems (Saah et al. 2019).

4.6. Transferability and operational recommendations

Our sampling optimisation framework is directly transferable to global arid and semi-arid ecosystems with heterogeneous woody vegetation across all cover density levels. This framework was specifically developed within the Malagasy spiny forest biome, a structurally coherent ecological unit; applying it to coarser geographic units would dilute vegetation heterogeneity and reduce threshold precision, whereas finer ecological subdivisions would improve stratification efficiency. Implementation requires: (i) Sentinel-1 + 2 combined data or AlphaEarth Foundations as predictors, (ii) photointerpretation via platforms like Collect Earth Online, and (iii) dispersed sampling representing all vegetation classes and cover types present in the landscape.

The quantitative thresholds identified (20.7–41.4 dispersed plots per 1,000 km^2 for optimal accuracy; 15.5 plots per 1,000 km^2 as minimal configuration) provide indicative reference benchmarks applicable to savannas, Sahelian zones, and xerophytic formations across continents, though independent recalibration is recommended for structurally distinct ecosystems or when using alternative modelling algorithms (Belgiu and Drăguţ 2016; Fassnacht et al. 2016). The comparative approach constitutes a standardised reproducible protocol applicable across environmental contexts. Compatibility with free Sentinel data and open-source photointerpretation platforms removes technological and financial barriers for developing countries. Managers can adapt sampling densities according to budget and local heterogeneity. For landscapes with greater structural complexity, we recommend increasing target sampling densities and adopting systematic stratified random sampling to ensure representation of all dominant ecosystem types (Fassnacht et al. 2014; Stehman and Foody 2019). Conversely, spatially homogeneous ecosystems are expected to reach performance saturation at lower densities (Mellor et al. 2015; Qu et al. 2024).

Finally, although AlphaEarth embeddings implicitly encode climatic signals, their individual contributions remain uninterpretable, making explicit precipitation predictors a valuable and transparent complement. Direct integration of freely available high-resolution precipitation data (e.g. CHIRPS, available in GEE at 5.5 km resolution) into Sentinel-based configurations therefore represents a straightforward improvement recommended for future applications, particularly where strong precipitation gradients drive vegetation heterogeneity.

5. Conclusion

This study demonstrates that very high-resolution photointerpretation provides a robust operational alternative to field inventories for calibrating woody cover models in arid ecosystems. Models calibrated with 505 photo-interpreted plots achieved substantially higher accuracy ($R^2 = 0.88$) compared to 41 field plots ($R^2 = 0.46–0.66$), confirming superiority of spatial representativeness over point-level precision.

These findings advance prior work in three directions. Our empirically derived thresholds (20.7–41.4 dispersed plots/1,000 km^2) align with saturation patterns reported by Mellor et al. (2015) beyond 200–300 samples for land cover mapping and by Qu et al. (2024) from approximately 150 samples for soil property prediction, while providing spatially normalised, biome-specific benchmarks absent from earlier studies. Our demonstration that spatial representativeness outweighs measurement precision extends the conclusions of Tsendbazar et al. (2021)—who showed this principle at the validation stage—to the calibration stage itself. Finally, the convergence of Sentinel-1 + 2 and AlphaEarth under adequate sampling reinforces Foody (2009) and Mellor et al. (2015), now demonstrated across foundation model embeddings.

Collectively, these results make six methodological contributions: (1) the first quantitative framework explicitly isolating calibration source, spatial configuration, and sampling density; (2) empirical evidence that spatial representativeness outweighs point-level precision; (3) a systematic cross-sensor comparison across four predictor sets; (4) operational sampling-density thresholds expressed as plots per 1,000 km^2; (5) demonstration that dispersed sampling achieves equivalent accuracy with half the samples of clustered designs; and (6) rigorous spatially blocked cross-validation ensuring reliable and generalisable performance estimates.

The identified thresholds provide directly transferable operational benchmarks for global arid ecosystems, implementable with Sentinel-1 + 2 or AlphaEarth predictors, dispersed photointerpretation via Collect Earth Online, and representative sampling across vegetation gradients. Compared to globally trained products such as Dynamic World ($R^2 = 0.70$ on the same independent field plots; Ramalason et al. 2025b), locally calibrated models with adequate photointerpretation sampling ($R^2 = 0.88$) demonstrate the substantial benefit of regional calibration, consistent with the growing evidence that spatially representative training data matter more than model complexity (Foody 2009; Tsendbazar et al. 2021). Principal limitations—single-biome derivation, dependence on Random Forest, and absence of formal inter-observer reliability assessment—define clear future priorities: threshold validation across biomes, evaluation of alternative algorithms, integration of high-resolution precipitation data (CHIRPS), and exploration of multi-temporal and drone-based acquisition.

Acknowledgements

This study was supported by the Académie de Recherche et de l'Enseignement Supérieur (ARES-CCD, Belgium) through the University of Liège mobility programme. The research was also conducted within the AGRIFO project (Sustainable Management and Restoration of Xerophilous Shrublands in southwestern Madagascar), also funded by ARES-CCD. We thank Dr. Ben Sparrow from Australian Terrestrial Ecosystem Research Network protocol for his support with the design and setup of the field inventory.

Author contributions

CRediT: **Felana Nantenaina Ramalason:** Conceptualization, Formal analysis, Investigation, Methodology, Software, Validation, Visualization, Writing – original draft, Writing – review & editing; **Olivia Lovanirina Rakotondrasoa:** Conceptualization, Funding acquisition, Project administration, Supervision; **Arthur Vander Linden:** Data curation, Formal analysis, Investigation, Methodology, Writing – review & editing; **Guillaume Renard:** Data curation, Formal analysis, Investigation, Methodology, Writing – review & editing; **Josoa R. Randriamalala:** Writing – review & editing; **Nicolas J. Vereecken:** Writing – review & editing; **Aina Razakamiaramanana:** Writing – review & editing; **Mbolatiana F. Ranaivoharivelo:** Writing – review & editing; **Sismondi Raholinarivo:** Writing – review & editing; **Harifidy Rakoto Ratsimba:** Methodology, Project administration, Resources; **Jan Bogaert:** Writing – review & editing; **Jean-François Bastin:** Conceptualization, Funding acquisition, Methodology, Project administration, Supervision, Validation, Writing – review & editing.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This research was supported by the Académie de recherche et d'enseignement supérieur.

ORCID

Felana Nantenaina Ramalason 0000-0003-1487-4027
Olivia Lovanirina Rakotondrasoa 0009-0009-8752-7406
Arthur Vander Linden 0000-0001-5760-9608
Guillaume Renard 0009-0006-1669-9495
Josoa R. Randriamalala 0000-0001-9313-0709
Nicolas J. Vereecken 0000-0002-8858-4623
Jan Bogaert 0000-0003-4465-6442
Jean-François Bastin 0000-0003-2602-7247

Data availability statement

The data and code supporting the findings of this study are openly available in Zenodo under a Creative Commons Attribution 4.0 International (CC-BY 4.0) license at https://doi.org/10.5281/zenodo.18235471. The dataset includes raw data, processed data, analysis code, results, and a readme file necessary to replicate all study findings reported in this article.

Data citation

Ramalason, F.N., Rakotondrasoa, O., Vander Linden, A., Renard, G., Randriamalala, J., Vereecken, N., Razakamiaramanana, A., Ranaivoharivelo, M.F., Raholinarivo, S., Rakoto Ratsimba, H., Bogaert, J. and Bastin, J.-F. (2026) *The importance of the sampling design in mapping woody cover in arid ecosystems* [Data set]. Zenodo. https://doi.org/10.5281/zenodo.18235471.

References

Abel, C., S. Horion, T. Tagesson, M. Brandt, and R. Fensholt. 2019. "Towards Improved Remote Sensing Based Monitoring of Dryland Ecosystem Functioning Using Sequential Linear Regression Slopes (Sergs)." *Remote Sensing of Environment* 224: 317–332. https://doi.org/10.1016/j.rse.2019.02.010.

Acharki, S., P. L. Frison, M. Amharref, H. Khoj, and S. Bernoussi. 2021. "Complémentarité Des Images Optiques SENTINEL-2 Avec Les Images Radar SENTINEL-1 Et ALOS-PALSAR-2 Pour La Cartographie De La Couverture végétale: Application à Une Aire protégée Et Ses Environs Au Nord-Ouest Du Maroc Via Trois Algorithmes D'Apprentissage Automatique." *RFPT* 223: 143–158. https://doi.org/10.52638/rfpt.2021.599.

Bastin, J.-F., N. Berrahmouni, A. Grainger, D. Maniatis, D. Mollicone, R. Moore, C. Patriarca, N. Picard, B. Sparrow, and E. M. Abraham. 2017. "The Extent of Forest in Dryland Biomes." *Science (New York, N.Y.)* 356: 635–638. https://doi.org/10.1126/science.aam6527.

Bastin, J.-F., Y. Finegold, C. Garcia, D. Mollicone, M. Rezende, D. Routh, C. M. Zohner, and T. W. Crowther. 2019. "The Global Tree Restoration Potential." *Science (New York, N.Y.)* 365: 76–79. https://doi.org/10.1126/science.aax0848.

Belgiu, M., and L. Drăguţ. 2016. "Random Forest in Remote Sensing: a Review of Applications and Future Directions." *ISPRS Journal of Photogrammetry and Remote Sensing* 114: 24–31. https://doi.org/10.1016/j.isprsjprs.2016.01.011.

Bey, A., A. Sánchez-Paus Díaz, D. Maniatis, G. Marchi, D. Mollicone, S. Ricci, J.-F. Bastin, et al. 2016. "Collect Earth: Land Use and Land Cover Assessment through Augmented Visual Interpretation." *Remote Sensing* 8: 807. https://doi.org/10.3390/rs8100807.

Brandt, M., C. J. Tucker, A. Kariryaa, K. Rasmussen, C. Abel, J. Small, J. Chave, et al. 2020. "An Unexpectedly Large Count of Trees in the West African Sahara and Sahel." *Nature (London, United Kingdom)* 587: 78–82. https://doi.org/10.1038/s41586-020-2824-5.

Brandt, M., P. Hiernaux, T. Tagesson, A. Verger, K. Rasmussen, A. A. Diouf, C. Mbow, E. Mougin, and R. Fensholt. 2016. "Woody plant cover estimation in drylands from Earth Observation based seasonal metrics." Remote Sensing of Environment 172: 28–38. https://doi.org/10.1016/j.rse.2015.10.036.

Breiman, L. 2001. "Random Forests." *Machine Learning* 45: 5–32. https://doi.org/10.1023/A:1010933404324.

Brown, C. F., S. P. Brumby, B. Guzder-Williams, T. Birch, S. B. Hyde, J. Mazzariello, W. Czerwinski, et al. 2022. "Dynamic World, Near Real-Time Global 10 M Land Use Land Cover Mapping." *Scientific Data* 9: 251. https://doi.org/10.1038/s41597-022-01307-4.

Brown, C. F., Kazmierski, M. R., Pasquarella, V. J., Rucklidge, W. J., Samsikova, M., Zhang, C., Shelhamer, E., et al. 2025. AlphaEarth Foundations: An embedding field model for accurate and efficient global mapping from sparse label data https://doi.org/10.48550/ARXIV.2507.22291.

Colditz, R. R., M. Schmidt, C. Conrad, M. C. Hansen, and S. Dech. 2011. "Land Cover Classification with Coarse Spatial Resolution Data to Derive Continuous and Discrete Maps for Complex Regions." *Remote Sensing of Environment* 115: 3264–3275. https://doi.org/10.1016/j.rse.2011.07.010.

Crowther, T. W., H. B. Glick, K. R. Covey, C. Bettigole, D. S. Maynard, S. M. Thomas, J. R. Smith, et al. 2015. "Mapping Tree Density at a Global Scale." *Nature (London, United Kingdom)* 525: 201–205. https://doi.org/10.1038/nature14967.

D'Odorico, P., A. Bhattachan, K. F. Davis, S. Ravi, and C. W. Runyan. 2013. "Global Desertification: Drivers and Feedbacks." *Advances in Water Resources* 51: 326–344. https://doi.org/10.1016/j.advwatres.2012.01.013.

De Marzo, T., D. Pflugmacher, M. Baumann, E. F. Lambin, I. Gasparri, and T. Kuemmerle. 2021. "Characterizing Forest Disturbances Across the Argentine Dry Chaco Based on Landsat Time Series." *International Journal of Applied Earth Observation and Geoinformation* 98: 102310. https://doi.org/10.1016/j.jag.2021.102310.

Dostálová, A., M. Doubková, D. Sabel, B. Bauer-Marschallinger, and W. Wagner. 2014. "Seven Years of Advanced Synthetic Aperture Radar (ASAR) Global Monitoring (GM) of Surface Soil Moisture over Africa." *Remote Sensing* 6: 7683–7707. https://doi.org/10.3390/rs6087683.

Drusch, M., U. Del Bello, S. Carlier, O. Colin, V. Fernandez, F. Gascon, B. Hoersch, et al. 2012. "Sentinel-2: ESA's Optical High-Resolution Mission for GMES Operational Services." *Remote Sensing of Environment* 120: 25–36. https://doi.org/10.1016/j.rse.2011.11.026.

Eldridge, D. J., M. A. Bowker, F. T. Maestre, E. Roger, J. F. Reynolds, and W. G. Whitford. 2011. "Impacts of Shrub Encroachment on Ecosystem Structure and Functioning: Towards a Global Synthesis." *Ecology letters* 14: 709–722. https://doi.org/10.1111/j.1461-0248.2011.01630.x.

Fassnacht, F. E., F. Hartig, H. Latifi, C. Berger, J. Hernández, P. Corvalán, and B. Koch. 2014. "Importance of Sample Size, Data Type and Prediction Method for Remote Sensing-Based Estimations of Aboveground Forest Biomass." *Remote Sensing of Environment* 154: 102–114. https://doi.org/10.1016/j.rse.2014.07.028.

Fassnacht, F. E., H. Latifi, K. Stereńczak, A. Modzelewska, M. Lefsky, L. T. Waser, C. Straub, and A. Ghosh. 2016. "Review of Studies on Tree Species Classification from Remotely Sensed Data." *Remote Sensing of Environment* 186: 64–87. https://doi.org/10.1016/j.rse.2016.08.013.

Feldman, A. F., S. Reed, C. Amaral, A. Babst-Kostecka, F. Babst, J. Biederman, and C. Devine, et al. 2024. "Adaptation and Response in Drylands (ARID): Community Insights for Scoping a NASA Terrestrial Ecology Field Campaign in Drylands." *Earth's Future* 12(9): e2024EF004811. https://doi.org/10.1029/2024ef004811.

Feng, S., and Q. Fu. 2013. "Expansion of Global Drylands under a Warming Climate." *Atmospheric Chemistry and Physics* 13: 10081–10094. https://doi.org/10.5194/acp-13-10081-2013.

Foody, G. M. 2009. "Classification Accuracy Comparison: Hypothesis Tests and the Use of Confidence Intervals in Evaluations of Difference, Equivalence and Non-Inferiority." *Remote Sensing of Environment* 113: 1658–1663. https://doi.org/10.1016/j.rse.2009.03.014.

Francis, R. J., R. T. Kingsford, K. Moseby, J. Read, R. Pedler, A. Fisher, J. McCann, and R. West. 2023. "Tracking Landscape Scale Vegetation Change in the Arid Zone By Integrating Ground, Drone and Satellite Data." *Remote Sensing in Ecology and Conservation* 10: 374–387. https://doi.org/10.1002/rse2.375.

Frison, P.-L., L. Jarlan, E. Mougin, P. Hiernaux, A. Dezetter, A. Alhassane, and J.-P. Rudant. 2018. "Forest Biomass Estimation in Tanzanian Miombo Woodlands Using Sentinel-1 and Sentinel-2 Data." *Remote Sensing* 10: 1439. https://doi.org/10.3390/rs10091439.

Funk, C., P. Peterson, M. Landsfeld, D. Pedreros, J. Verdin, S. Shukla, G. Husak, et al. 2015. "The Climate Hazards Infrared Precipitation with stations—a New Environmental Record for Monitoring Extremes." *Scientific Data* 2: 150066. https://doi.org/10.1038/sdata.2015.66.

Goldblatt, R., A. Rivera Ballesteros, and J. Burney. 2017. "High Spatial Resolution Visual Band Imagery Outperforms Medium Resolution Spectral Imagery for Ecosystem Assessment in the Semi-Arid Brazilian Sert~ Ao." *Remote Sensing* 9: 1336. https://doi.org/10.3390/rs9121336.

Guida-Johnson, B., P. E. Villagra, L. M. Alvarez, F. Rojas, and J. A. Alvarez. 2021. "Finding Woodlands in Drylands: Bases for the Monitoring of Xeric Open Forests in a Cloud Computing Platform." *Remote Sensing Applications: Society and Environment* 22: 100528. https://doi.org/10.1016/j.rsase.2021.100528.

Hänke, H., J. Barkmann, C. Coral, E. Enfors Kaustky, and R. Marggraf. 2017. "Social-Ecological Traps Hinder Rural Development in Southwestern Madagascar." *E&S* 22: art42. https://doi.org/10.5751/ES-09130-220142.

Hawkins, D. M. 2004. "The Problem of Overfitting." *Journal of Chemical Information and Computer Sciences* 44(1): 1–12. https://doi.org/10.1021/ci0342472.

Hersbach, H., B. Bell, P. Berrisford, S. Hirahara, A. Horányi, J. Muñoz-Sabater, J. Nicolas, et al. 2020. "The ERA5 Global Reanalysis." *Quarterly Journal of the Royal Meteorological Society* 146: 1999–2049. https://doi.org/10.1002/qj.3803.

Karakizi, C., O. Gounari, E. Sofikiti, G. Begkos, K. Karantzalos, and E. Symeonakis. 2023. Assessing the Contribution of Optical and Sar Data for Fractional Savannah Woody Vegetation Mapping. In *IGARSS 2023 - 2023 IEEE International Geoscience and Remote Sensing Symposium*. 3118–3121, Pasadena, CA, USA: IEEE International Geoscience and Remote Sensing Symposium, IEEE, Presented at the IGARSS 2023. https://doi.org/10.1109/IGARSS52108.2023.10282969.

Karl, J. W., S. E. McCord, and B. C. Hadley. 2017. "A Comparison of Cover Calculation Techniques for Relating Point-Intercept Vegetation Sampling to Remote Sensing Imagery." *Ecological Indicators* 73: 156–165. https://doi.org/10.1016/j.ecolind.2016.09.034.

IPCC Lee, H. and J. Romero, eds. 2023. Climate Change 2023: Synthesis Report. Contribution of Working Groups I, II and III to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change [Core Writing Team].Geneva, Switzerland: IPCC. Intergovernmental Panel on Climate Change (IPCC). https://doi.org/10.59327/IPCC/AR6-9789291691647.

McRoberts, R. E., Q. Chen, D. D. Gormanson, and B. F. Walters. 2018. "Probability-And Model-Based Approaches to Inference for Proportion Forest Using Satellite Imagery as Auxiliary Data." *Remote sensing of environment* 204: 642–649. https://doi.org/10.1016/j.rse.2017.09.032.

Mellor, A., S. Boukir, A. Haywood, and S. Jones. 2015. "Exploring Issues of Training Data Imbalance and Mislabelling on Random Forest Performance for Large Area Land Cover Classification Using the Ensemble Margin." *ISPRS Journal of Photogrammetry and Remote Sensing* 105: 155–168. https://doi.org/10.1016/j.isprsjprs.2015.03.014.

Meyer, H., Ludwig, M, Milà,, C., Linnenbrink, J., and Schumacher, F. 2024. The CAST package for training and assessment of spatial prediction models in R [WWW Document]. arXiv.org. URL (accessed 10.28.25). https://doi.org/10.48550/arXiv.2404.06978.

Nuzzo, R. L. 2019. "Histograms: a Useful Data Analysis Visualization." *PM&R* 11: 309–312. https://doi.org/10.1002/pmrj.12145.

Olofsson, P., G. M. Foody, M. Herold, S. V. Stehman, C. E. Woodcock, and M. A. Wulder. 2014. "Good Practices for Estimating Area and Assessing Accuracy of Land Change." *Remote Sensing of Environment* 148: 42–57. https://doi.org/10.1016/j.rse.2014.02.015.

Pengra, B. W., S. V. Stehman, J. A. Horton, D. J. Dockter, T. A. Schroeder, Z. Yang, W. B. Cohen, S. P. Healey, and T. R. Loveland. 2007. "Quality Control and Assessment of Interpreter Consistency of Annual Land Cover Reference Data in an Operational National Monitoring Program." *ISPRS journal of photogrammetry and remote sensing* 128: 299–306. https://doi.org/10.1016/j.isprsjprs.2017.04.010.

Pérez-Luque, A. J., M. E. Ramos-Font, M. J. Tognetti Barbieri, C. Tarragona Pérez, G. Calvo Renta, and A. B. Robles Cruz. 2022. "Vegetation Cover Estimation in Semi-Arid Shrublands after Prescribed Burning: Field-Ground and Drone Image Comparison." *Drones* 6: 370. https://doi.org/10.3390/drones6110370.

Ploton, P., F. Mortier, M. Réjou-Méchain, N. Barbier, N. Picard, V. Rossi, C. Dormann, G. Cornu, G. Viennois, and N. Bayol. 2020. "Spatial Validation Reveals Poor Predictive Performance of Large-Scale Ecological Mapping Models." *Nature communications* 11: 4540. https://doi.org/10.1038/s41467-020-18321-y.

Pouliot, D., R. Latifovic, J. Pasher, and J. Duffe. 2018. "Landsat Super-Resolution Enhancement Using Convolution Neural Networks and Sentinel-2 for Training." *Remote Sensing* 10: 394. https://doi.org/10.3390/rs10030394.

Probst, P., M. N. Wright, and A. Boulesteix. 2019. "Hyperparameters and Tuning Strategies for Random Forest." *WIREs Data Min & Knowl* 9(3): e1301. https://doi.org/10.1002/widm.1301.

Provencher-Nolet, L., M. Bernier, and E. Lévesque. 2014. Short Term Change Detection in Tundra Vegetation Near Umiujaq, Subarctic Quebec, Canada In *2014 IEEE Geoscience and Remote Sensing Symposium*4668–4670. IEEE Geoscience and Remote Sensing Symposium, Presented at the 2014. https://doi.org/10.1109/IGARSS.2014.6947534.

Qu, L., H. Lu, Z. Tian, J. M. Schoorl, B. Huang, Y.Liang, D. Qiu , and Y. Liang. 2024. "Spatial Prediction of Soil Sand Content at Various Sampling Density Based on Geostatistical and Machine Learning Algorithms in Plain Areas." *CATENA* 234: 107572. https://doi.org/10.1016/j.catena.2023.107572.

Ramalason, F. N., O. L. Rakotondrasoa, and J.-F. Bastin. 2025a. Recension Des Travaux antérieurs Sur Les méthodes utilisées Dans La Cartographie Des écosystèmes Forestiers Des Zones Arides: Cas Du Sud-Ouest De Madagascar. In *LES CAHIERS DE L'ED GRND. Presented at the Colloque de l'ED GRND 2023*, ED GRND, Antananarivo, Madagascar.

Ramalason, F. N., O. L. Rakotondrasoa, A. V. Linden, G. Renard, H. R. Ratsimba, J. Bogaert, and J.-F. Bastin. 2025b. "Potential of Global Vegetation Maps in Capturing Xerophytic Vegetation Cover: Insights from Madagascar'S Arid Ecosystems." *Journal of Arid Environments* 229: 105406. https://doi.org/10.1016/j.jaridenv.2025.105406.

Randriamalala, J. R., H. O. Radosy, M. Ramanakoto, H. Razafindrahanta, J.-M. Ravoninjatovo, R. S. Haingomanantsoa, and T. Ramananantoandro. 2022. "Allometric Models to Predict the Individual Aboveground Biomass of Shrubs of Malagasy Xerophytic Thickets." *Journal of Arid Environments* 202: 104751. https://doi.org/10.1016/j.jaridenv.2022.104751.

Reiche, J., R. Verhoeven, J. Verbesselt, E. Hamunyela, N. Wielaard, and M. Herold. 2018. "Characterizing Tropical Forest Cover Loss Using Dense Sentinel-1 Data and Active Fire Alerts." *Remote Sensing* 10: 777. https://doi.org/10.3390/rs10050777.

Reynolds, J. F., D. M. Stafford Smith, E. F. Lambin, B. L. Turner, M. Mortimore, S. P. Batterbury, T. E. Downing, H. Dowlatabadi, R. J. Fernández, and J. E. Herrick. 2007. "Global Desertification: Building a Science for Dryland Development." *Science (New York, N.Y.)* 316: 847–851. https://doi.org/10.1126/science.1131634.

Saah, D., G. Johnson, B. Ashmall, G. Tondapu, K. Tenneson, M. Patterson, A. Poortinga, et al. 2019. "Collect Earth: an Online Tool for Systematic Reference Data Collection in Land Cover and Use Applications." *Environmental Modelling & Software* 118: 166–171. https://doi.org/10.1016/j.envsoft.2019.05.004.

Sengani, D., A. Ramoelo, and E. Archer. 2023. "A Review of Fusion Framework Using Optical Sensors and Synthetic Aperture Radar Imagery to Detect and Map Land Degradation and Sustainable Land Management in the Semi-Arid Regions." *Geocarto International* 38: 2278325. https://doi.org/10.1080/10106049.2023.2278325.

Stehman, S. V., and G. M. Foody. 2019. "Key Issues in Rigorous Accuracy Assessment of Land Cover Products." *Remote Sensing of Environment* 231: 111199. https://doi.org/10.1016/j.rse.2019.05.018.

Stock, A. 2025. "Choosing Blocks for Spatial Cross-Validation: Lessons from a Marine Remote Sensing Case Study." *Frontiers in Remote Sensing*6: 1531097. https://doi.org/10.3389/frsen.2025.1531097.

Symeonakis, E., A. Korkofigkas, T. Higginbottom, J. Boyd, E. Arnau-Rosalen, G. Stamou, and K. Karantzalos. 2022. Towards a Deep Learning Fractional Woody Vegetation Cover Monitoring Framework. In *IGARSS 2022 - 2022 IEEE International Geoscience and Remote Sensing Symposium*5905–5908, Kuala Lumpur, Malaysia: IEEE, Presented at the IGARSS 2022. https://doi.org/10.1109/IGARSS46834.2022.9883347.

Torres, R., P. Snoeij, D. Geudtner, D. Bibby, M. Davidson, E. Attema, P. Potin, et al. 2012. "GMES Sentinel-1 Mission." *Remote Sensing of Environment* 120: 9–24. https://doi.org/10.1016/j.rse.2011.05.028.

Trabucco, A., and R. Zomer. 2019. *Global Aridity Index and Potential Evapotranspiration (ET0) Climate Database* v2. https://doi.org/10.6084/M9.FIGSHARE.7504448.V3.

Tsendbazar, N., M. Herold, L. Li, A. Tarko, S. de Bruin, D. Masiliunas, M. Lesiv, et al. 2021. "Towards Operational Validation of Annual Global Land Cover Maps." *Remote Sensing of Environment* 266: 112686. https://doi.org/10.1016/j.rse.2021.112686.

UNCCD. 2024. The Global Threat of Drying Lands: Regional and global aridity trends and future projections. A Report of the Science-Policy Interface.

Valavi, R., J. Elith, J. Lahoz-Monfort, and G. Guillera-Arroita. 2018. "blockCV: an R Package for Generating Spatially or Environmentally Separated Folds for k-fold Cross-Validation of Species Distribution Models." *Methods in Ecology and Evolution* 10: 225–232. https://doi.org/10.1111/2041-210X.13107.

Valjarević, A., A. Šiljeg, S. Šiljeg, F. Vujović, and A. Sahay. 2025. "GIS-based Water Stress Analysis in North African Drylands." *Journal of Arid Environments* 230: 105427. https://doi.org/10.1016/j.jaridenv.2025.105427.

Valjarević, A., T. Djekić, V. Stevanović, R. Ivanović, and B. Jandziković. 2018. "GIS Numerical and Remote Sensing Analyses of Forest Changes in the Toplica Region for the Period of 1953-2013." *Applied Geography* 92: 131–139. https://doi.org/10.1016/j.apgeog.2018.01.016.

Van Passel, J., W. De Keersmaecker, and B. Somers. 2020. “Monitoring Woody Cover Dynamics in Tropical Dry Forest Ecosystems Using sentinel-2 Satellite Imagery.” *Remote Sensing* 12: 1276. https://doi.org/10.3390/rs12081276.

Wang, G., G. Gertner, and A. B. Anderson. 2005. “Sampling Design and Uncertainty Based on Spatial Variability of Spectral Variables for Mapping Vegetation Cover.” *International Journal of Remote Sensing* 26: 3255–3274. https://doi.org/10.1080/01431160500114748.

White, A., B. Sparrow, E. Leitch, J. Foulkes, R. Flitton, A. J. Lowe, and S. Caddy-Retalic. 2012. *AusPlots Rangelands survey protocols manual*. University of Adelaide Press.

Yamazaki, D., D. Ikeshima, J. Sosa, P. D. Bates, G. H. Allen, and T. M. Pavelsky. 2019. “MERIT Hydro: a High-Resolution Global Hydrography Map Based on Latest Topography Dataset.” *Water Resources Research* 55: 5053–5073. https://doi.org/10.1029/2019WR024873.

Zhang, X., Q. Pan, Y. Zhao, and Q Huang. 2005. Research on spatial deviation analysis model of land-use change. In Presented at the MIPPR 2005 Image Analysis Techniques D. Li and H. Ma, eds. 60440C. Wuhan, China. https://doi.org/10.1117/12.652320.