Journal of Applied Informatics and Computing (JAIC)
Vol.10, No.2, April 2026, pp. 1520~1529
e-ISSN: 2548-6861

Design and Evaluation of an Optimized Random Forest Classification Framework for Tropical Forest Degradation Detection

Frey Sylvestre [1,3]*, Simboni Simboni Tege [2], Mabela Makengo Ronstand [1,3], Mbuyi Mukendi Eugene [1], Anzola Kibamba Nestor [1], Kafunda Katalay Pierre [1]
[1] Department of Mathematics, Statistics and Computer Science, University of Kinshasa, Kinshasa, D.R. Congo
[2] Department of Computer Management, Higher Pedagogical Institute of Isiro, Isiro, D.R. Congo
[3] Research Center for Mathematics Education (CREM)
sylvestre.frey@unikin.ac.cd [1], tege.simboni1@gmail.com [2], rostin.mabela@unikin.ac.cd [3], nestor.anzola@unikin.ac.cd[5], pierre.kafunda@unikin.ac.cd [6]

Article Info

Article history:

Received 2026-02-14
Revised 2026-03-02
Accepted 2026-04-10

Keyword:

Random Forest,
Tropical forest monitoring,
Automated classification,
Change detection, Deforestation,
Machine learning,
Landsat 8,
Congo Basin.

ABSTRACT

Tropical forest degradation monitoring in resource-constrained environments requires classification approaches that balance predictive accuracy with computational efficiency. While deep learning methods often achieve high classification performance, their reliance on large labeled datasets and GPU-based infrastructure limits operational deployment in many conservation contexts. This study proposes and evaluates a systematically optimized Random Forest classification framework for tropical forest degradation detection in the Congo Basin using Landsat 8 multispectral imagery. A multi-class model distinguishing intact forest, degraded forest, and non-forest was developed using six spectral bands and four ecologically motivated vegetation indices (NDVI, EVI, NDMI, and NBR). Ground truth data consisted of 450 stratified samples (180 intact forest, 150 degraded forest, 120 non-forest; imbalance ratio 1.5) validated through field surveys and high-resolution imagery collected during the 2022 dry season. Class imbalance was mitigated through balanced class weighting during training. Model optimization was achieved through controlled hyperparameter tuning with five-fold cross-validation, targeting improved generalization while preserving computational efficiency. The optimized Random Forest achieved an overall accuracy of 89.3% (Kappa = 0.834) on an independent test set, significantly outperforming CART and SVM baselines while maintaining a training time under two minutes on standard CPU hardware. Class-specific F1-scores ranged from 0.88 to 0.91, indicating balanced performance across all land-cover classes. Processing the full 7500 km² study area required approximately 14 minutes without GPU acceleration. These results demonstrate that systematically optimized ensemble learning provides a robust, interpretable, and operationally deployable solution for tropical forest degradation monitoring, particularly suited to conservation agencies operating under limited computational and financial resources.

I. INTRODUCTION

Recent developments in artificial intelligence have transformed environmental monitoring at continental scales. In remote sensing applications, machine learning algorithms now enable automated analysis across millions of square kilometers (Maxwell et al., 2018). Deep learning approaches, particularly Convolutional Neural Networks (CNNs), have demonstrated classification accuracies exceeding 90% on satellite imagery (Zhang et al., 2021). However, these sophisticated techniques demand substantial computational

infrastructure: GPU clusters, labeled datasets with more than 10,000 annotated samples, and training durations often surpassing 24 hours requirements that remain out of reach for many operational conservation programs working under resource constraints (Campos-Taberner et al., 2020).

The Congo Basin contains the world's second-largest tropical rainforest system, storing approximately 30 billion tonnes of carbon across six million square kilometers while supporting exceptional biodiversity (Lewis et al., 2009). The Democratic Republic of Congo, which encompasses 60% of this critical ecosystem, faces accelerating deforestation with annual forest cover loss reaching 490,000 hectares, representing 0.23% of total forest area (FAO, 2020; Potapov et al., 2012). This degradation threatens not only regional climate stability and biodiversity conservation, but also the livelihoods of 40 million people whose communities depend directly on forest resources (De Wasseige et al., 2012).

Traditional forest monitoring approaches rely primarily on ground patrol networks, manual satellite imagery interpretation, and periodic field surveys (Fuller, 2006). These conventional techniques suffer from fundamental limitations that severely restrict their effectiveness across vast tropical landscapes. Ground patrols, while providing high-quality localized data, achieve coverage of less than 2% of protected areas annually due to logistical constraints, difficult terrain, and limited human resources (ICCN, 2012). Manual satellite image interpretation introduces delays of typically three to six months between image acquisition and detection, rendering interventions reactive rather than preventive (Herold & Skutsch, 2011).

Several global automated forest monitoring systems have emerged to address these limitations. Global Forest Watch (Hansen et al., 2013) provides annual forest loss maps via decision trees; FORMA (Hammer et al., 2014) detects changes in near-real-time using MODIS 500m imagery; and GLAD Alerts (Reiche et al., 2018) combine Landsat and Sentinel-2 for weekly detection at 30m resolution. While effective at pan-tropical scales, these platforms present three critical limitations for local conservation applications. They optimize for average global performance at the expense of precision in specific ecosystems such as Congo Basin peatlands. All detected changes receive uniform treatment without consideration of ecological priorities or intervention feasibility. Finally, their use requires advanced geomatics expertise, limiting adoption by local ranger teams and park managers (Tyukavina et al., 2018).

From a computer science and systems engineering perspective, forest degradation detection should be understood as a large-scale multi-class classification problem applied to high-dimensional, heterogeneous spatial data. The central challenge lies in designing machine learning models that simultaneously achieve high predictive robustness, strong spatial generalization capacity, and computational efficiency, while respecting strict operational constraints: limited hardware resources, extensive spatial coverage, and frequent temporal updates (Maxwell et al., 2018; Belgiu & Drăguţ, 2016). In this context, algorithmic optimization serves a purpose beyond improving statistical performance—it directly conditions the operational feasibility of large-scale surveillance systems (Breiman, 2001; Rodriguez-Galiano et al., 2012).

This study addresses these challenges through the development of an optimized Random Forest classification framework specifically adapted to operational constraints in the Congo Basin. The proposed approach balances predictive accuracy with computational efficiency to enable deployment on standard computing hardware available to local conservation agencies. The framework distinguishes three critical forest states —intact forest, degraded forest, and non-forest areas—using Landsat 8 multispectral imagery combined with derived vegetation indices.

The novelty of this study does not reside in algorithmic innovation—Random Forest is a well-established method—but rather in three original and reproducible contributions. First, we propose a systematic hyperparameter optimization pipeline (RandomizedSearchCV with stratified 5-fold cross-validation) applied to a swamp forest context not previously studied with this method. Second, we present an ecologically motivated selection of four complementary spectral indices (NDVI, EVI, NDMI, NBR), validated through feature importance analysis within the specific hydrological regime of Tumba-Lediima. Third, we demonstrate that this approach achieves classification performance comparable to deep learning methods (accuracy >89%) while remaining deployable on standard CPU hardware training times under two minutes versus hours of GPU processing thereby establishing a reproducible benchmark for conservation agencies operating without GPU infrastructure.

The remainder of this paper proceeds as follows. Section II reviews relevant literature on machine learning applications in forest monitoring and positions our approach within current methodological debates. Section III details our experimental design, including study area selection, data acquisition and preprocessing, feature engineering, and model training procedures. Section IV presents classification results, comparative performance analysis, and computational efficiency metrics. Section V discusses findings in relation to operational deployment requirements and identifies directions for future research. Section VI concludes with a synthesis of key contributions and their practical implications for conservation practice.

II. LITERATURE REVIEW

A. Machine Learning in Remote Sensing

Supervised learning approaches vary considerably in their suitability for forest classification tasks. Simple decision trees, while interpretable, tend to overfit on complex multispectral data (Friedl & Brodley, 1997). Support Vector Machines have shown promise in earlier studies but struggle with large training datasets-a significant limitation for continental-scale applications (Mountrakis et al., 2011). By contrast, ensemble methods such as Random Forest have

emerged as particularly well-suited to remote sensing applications, offering robust performance across diverse ecosystems (Belgiu & Drăguţ, 2016).

Random Forest operates by constructing multiple decision trees during training and outputting the mode of class predictions from individual trees. This ensemble approach provides several advantages for satellite image classification. Bootstrap aggregation (bagging) reduces overfitting compared to single decision trees. Out-of-bag samples enable internal validation without requiring separate test datasets. Feature importance metrics facilitate interpretation of which spectral bands or indices contribute most to classification accuracy (Breiman, 2001).

Deep learning methods, particularly Convolutional Neural Networks, have attracted significant attention in recent remote sensing literature (Zhang et al., 2021). These approaches can automatically learn hierarchical feature representations from raw imagery, potentially capturing complex spatial patterns that handcrafted features might miss. However, CNNs typically require thousands of labeled training samples to avoid overfitting a threshold difficult to achieve in field-based forest monitoring programs where ground truth collection proves time-consuming and expensive. Training deep networks also demands GPU infrastructure and may require days of computation time, even for moderately-sized study areas (Ma et al., 2019).

B. Tropical Forest Monitoring Challenges

Tropical forests present distinct classification challenges compared to temperate regions. Persistent cloud cover in equatorial zones reduces usable satellite imagery, particularly during wet seasons when acquisition frequencies may drop below monthly intervals (Asner, 2001). High species diversity creates complex spectral mixing within forest canopies, complicating efforts to distinguish degradation from natural heterogeneity. Seasonal flooding in lowland areas, such as the peatland forests of the Congo Basin, further introduces temporal variability in spectral signatures that complicates year-round monitoring (Dargie et al., 2017).

Forest degradation detection proves more challenging than outright deforestation detection. Clear-cutting produces dramatic spectral changes easily identified through simple vegetation index thresholds. Degradation manifests more subtly selective logging may remove only 10-15% of canopy cover, falling below detection limits of moderate-resolution sensors. Small-scale clearings for agriculture or charcoal production may occupy areas smaller than a single Landsat pixel (900 m^2), requiring sub-pixel classification approaches or higher-resolution imagery (Souza et al., 2013).

The Congo Basin specifically presents additional monitoring complications. Political instability in parts of the region restricts field access for ground truth collection. Limited technical capacity within local conservation agencies constrains adoption of sophisticated monitoring tools. Infrastructure limitations including unreliable electricity and limited internet bandwidth restrict deployment of computationally intensive classification approaches requiring cloud computing resources (Shapiro et al., 2016).

C. Operational Considerations for Conservation Applications

Forest monitoring systems, to prove operationally viable, must satisfy requirements beyond statistical accuracy. Detection latency directly impacts intervention effectiveness alerts generated months after degradation occurs provide limited actionable value. Processing efficiency determines whether systems can scale from single protected areas to entire landscapes. Model interpretability affects trust and adoption by conservation practitioners who may lack machine learning expertise but require confidence in automated detections (Reiche et al., 2018).

Cost-effectiveness represents a critical yet often overlooked consideration. Many conservation organizations in developing countries operate on annual budgets under $100,000 USD, rendering subscription fees for commercial satellite data or cloud processing prohibitive. Open-access Landsat imagery, despite its moderate 30m resolution, provides a sustainable data source for long-term monitoring programs. Similarly, algorithms that train efficiently on standard laptop computers prove more deployable than those requiring specialized GPU infrastructure (Woodcock et al., 2008).

Prior research has demonstrated successful applications of machine learning to forest monitoring. Rodriguez-Galiano et al. (2012) achieved 91% accuracy classifying Mediterranean forests using Random Forest with Landsat imagery, though their study area lacked the spectral complexity of tropical forests. Belgiu & Drăguţ (2016) reviewed Random Forest applications across diverse remote sensing tasks, concluding that the algorithm demonstrates particular robustness when training sample sizes prove limited a common situation in field-based validation programs. However, few studies have explicitly addressed the operational deployment challenges faced by conservation agencies in resource-constrained environments.

This study builds on prior work by developing an optimized Random Forest framework specifically designed for operational deployment in the Congo Basin. Unlike previous research focusing primarily on classification accuracy, we emphasize computational efficiency and deployment feasibility as co-equal objectives with statistical performance. Our approach demonstrates that careful algorithm optimization and feature engineering can achieve results comparable to more sophisticated methods while maintaining compatibility with standard computing hardware available to local conservation agencies.

III. METHODOLOGY

A. Study Area

Tumba-Lediima Nature Reserve occupies approximately 7500 km² in the Mai-Ndombe Province of western Democratic Republic of Congo (approximately 1°15'S–1°45'S, 18°30'E–19°00'E). The reserve lies within a low-elevation floodplain (300–350 m) receiving between 1,800 and 2,000 mm of annual rainfall, supporting one of the region's most extensive peatland complexes (Dargie et al., 2017). This ecosystem represents a critical carbon store, with peat deposits estimated to contain approximately 30.6 gigatonnes of carbon equivalent to 15 years of current U.S. fossil fuel emissions.

The reserve's swamp forest vegetation presents particular challenges for remote sensing classification. Seasonal flooding patterns, persistent cloud cover averaging 60–70% throughout the year, and spectral similarities between intact and degraded forest under varying moisture conditions complicate automated detection approaches. These characteristics make Tumba-Lediima a demanding test case for classification algorithms, and success here would suggest strong potential for transfer to other Congo Basin protected areas.

Deforestation pressures stem primarily from small-scale agriculture expansion and artisanal logging for construction materials and fuelwood. Unlike industrial logging operations that create large clearings easily detected from satellite imagery, degradation in Tumba-Lediima typically manifests as dispersed small clearings of 0.1 to 0.5 hectares distributed throughout the forest matrix. This spatial pattern requires classification approaches capable of detecting disturbances at or near the spatial resolution limit of Landsat imagery (30 m pixels).

B. Data Acquisition and Preprocessing

We acquired cloud-free Landsat 8 OLI (Operational Land Imager) scenes covering the study area from the U.S. Geological Survey archive for the June–August 2022 dry season. Scene selection targeted acquisition dates within a 10-day window to minimize phenological variation across the mosaic. The Landsat 8 sensor provides multispectral imagery in seven reflective bands spanning visible to shortwave infrared wavelengths (Table 1), with 30-meter spatial resolution suitable for detecting small-scale forest disturbances.

TABLE 1
LANDSAT 8 OLI SPECTRAL BANDS EMPLOYED IN CLASSIFICATION

Band	Wavelength (μm)	Spatial Resolution
Band 2 (Blue)	0.450-0.515	30 m
Band 3 (Green)	0.525-0.600	30 m
Band 4 (Red)	0.630-0.680	30 m
Band 5 (NIR)	0.845-0.885	30 m
Band 6 (SWIR1)	1.560-1.660	30 m
Band 7 (SWIR2)	2.100-2.300	30 m

Dark Object Subtraction (DOS) atmospheric correction was applied to all scenes (Chavez, 1996), and cloud contamination was addressed through the Landsat 8 quality assessment band. Following preprocessing, four spectral indices were computed (Table 2). NDVI quantifies vegetation density and vigor, while EVI improves sensitivity in high-biomass areas where NDVI saturates. NDMI captures canopy moisture content via NIR and SWIR1 contrast a particularly discriminating variable in the Tumba-Lediima peatland context where moisture gradients are a key structural factor. NBR, originally designed for burned area mapping, effectively highlights canopy damage and recent disturbance from selective cutting through NIR and SWIR2 contrasts. This combination was selected to ensure no two indices are spectrally redundant, a property confirmed by the feature importance analysis presented in Section IV.

TABLE 2
SPECTRAL VEGETATION INDICES CALCULATED FROM LANDSAT 8 BANDS

Index	Formula	Ecological Significance
NDVI	(NIR - Red) / (NIR + Red)	Vegetation vigor, biomass
EVI	2.5 × ((NIR - Red) / (NIR + 6×Red - 7.5×Blue + 1))	Enhanced sensitivity in high biomass areas
NDMI	(NIR - SWIR1) / (NIR + SWIR1)	Canopy moisture content
NBR	(NIR - SWIR2) / (NIR + SWIR2)	Recent disturbance, canopy damage

C. Ground Truth Data Collection and Class Distribution

We established 450 sampling points using a stratified random sampling design. Initial stratification was based on visual interpretation of high-resolution imagery (0.5 m WorldView-3 and 10 m Sentinel-2 composites). The final ground truth dataset comprised 180 intact forest points (>70% canopy closure, no recent disturbance), 150 degraded forest points (30–70% canopy closure, evidence of selective logging or small clearings), and 120 non-forest points (agricultural fields, open water, herbaceous wetlands, settlements).

This distribution (40% / 33% / 27%) reflects the study area's actual land cover composition while ensuring adequate representation of all classes. Although classes are not perfectly balanced, the imbalance ratio (max/min = 1.5) remains below the threshold commonly associated with significant classification bias in Random Forest (Japkowicz & Stephen, 2002). To further mitigate potential imbalance effects, the class_weight='balanced' parameter was applied during model training, ensuring proportional class contribution to the loss function. The resulting sampling density of approximately one point per 16.5 km² reflects the logistical constraints characteristic of field access in the Congo Basin. While lower than densities reported in more accessible study areas (Rodriguez-Galiano et al., 2012: one point per 2.1 km²), it remains consistent with studies conducted in similar field-constrained tropical environments

(Belgiu & Drăguț, 2016) and is sufficient given the spatial resolution of Landsat 8.

Field validation was conducted during July–August 2022, coinciding with the Landsat 8 image acquisition period and thereby minimizing temporal mismatch between ground observations and satellite data. Each point was visited by field teams equipped with GPS units (±3 m accuracy), documenting land cover type, canopy closure, dominant species, and evidence of disturbance. Where field access proved impossible (~15% of points), very-high-resolution commercial satellite imagery was substituted.

The 450 points were partitioned 70/30 using stratified sampling: 315 training samples and 135 test samples. The 315 training samples were used exclusively for 5-fold cross-validation during hyperparameter optimization; the 135 test samples remained strictly held out and were never exposed during training or tuning.

D. Feature Engineering and Selection

The initial feature space comprised six original Landsat 8 reflective bands plus four derived spectral indices, yielding ten predictor variables. Preliminary feature importance analysis via mean decrease in Gini impurity identified NDVI (32%), NDMI (24%), NIR band (18%), and NBR (12%) as the most discriminating variables, with cumulative importance exceeding 86%. We retained the full ten-feature set rather than applying aggressive dimensionality reduction: Random Forest's internal bootstrap sampling handles weakly predictive features with minimal computational overhead, and retaining interpretable spectral bands alongside indices facilitates validation by domain experts familiar with vegetation spectral signatures.

Analysis of predictor variable distributions across land cover classes revealed substantial overlap between intact and degraded forest in individual spectral bands. Multi-dimensional separation, however, proved more distinct supporting the use of ensemble methods capable of learning complex decision boundaries rather than simple threshold-based classifiers.

E. Model Development and Hyperparameter Optimization

We implemented Random Forest classification using scikit-learn 1.2 in Python 3.9 on a standard laptop (Intel Core i7-8550U, 16 GB RAM, no dedicated GPU). Hyperparameter optimization targeted three key parameters via 5-fold cross-validated RandomizedSearchCV. The n_estimators grid [100, 200, 300, 500] showed performance plateauing around 300 trees, with marginal gains below 0.5% beyond that threshold. For max_depth [10, 15, 20, None], unlimited depth caused overfitting with a training–validation gap exceeding 5%; a value of 20 was selected. For min_samples_leaf [5, 10, 20, 50], a value of 10 reduced overfitting while maintaining accuracy. The final optimized configuration n_estimators=300, max_depth=20, min_samples_leaf=10, max_features='sqrt', class_weight='balanced' achieved 89.3% cross-validation accuracy with a training time of 1.8 minutes on the specified hardware.

F. Baseline Model Comparison

To contextualize Random Forest performance, we implemented two baseline classifiers: Classification and Regression Trees (CART) and Support Vector Machines (SVM) with radial basis function (RBF) kernel. CART represents a simpler single-tree approach offering interpretability advantages but typically lower ensemble accuracy. SVM-RBF has demonstrated strong performance in prior remote sensing applications, particularly for high-dimensional data.

Both baseline models received identical input features and underwent analogous hyperparameter optimization through grid search cross-validation. This ensured fair comparison focused on algorithmic differences rather than tuning artifacts. CART optimization varied tree depth and minimum split samples. SVM tuning explored regularization parameter C and RBF kernel width gamma across logarithmic ranges.

G. Accuracy Assessment

Classification performance was evaluated using standard remote sensing accuracy metrics: overall accuracy, producer's accuracy (recall), user's accuracy (precision), F1-score, and Kappa coefficient. The confusion matrix structure enabled identification of class-specific errors, revealing whether misclassifications concentrated in particular class pairs notably potential confusion between intact and degraded forest. Kappa values above 0.80 indicate substantial agreement (Landis & Koch, 1977), providing a chance-corrected complement to overall accuracy that is especially informative under unequal class distributions.

Beyond statistical metrics, we recorded training time, full-area prediction time, and memory requirements as operational performance indicators. In resource-constrained conservation contexts, these computational efficiency benchmarks are equally important as classification accuracy for determining real-world deployment feasibility.

IV. RESULTS

A. Classification Performance

The optimized Random Forest model achieved an overall accuracy of 89.3% (Kappa: 0.834) on the held-out test dataset (n=135). This places the model within the range reported for comparable tropical forest classification studies while maintaining computational efficiency suitable for operational deployment.

TABLE 3
COMPARATIVE ACCURACY METRICS ACROSS CLASSIFICATION ALGORITHMS

Model	Overall Accuracy	Kappa	Training Time	Prediction Time (7500 km²)
CART	83.1%	0.741	12 seconds	8 minutes
SVM (RBF)	85.7%	0.780	8 minutes	22 minutes
Random Forest (optimized)	89.3%	0.834	1.8 minutes	14 minutes

The Random Forest (optimised) achieves the highest overall accuracy and Kappa coefficient while maintaining a training time of 1.8 minutes significantly lower than SVM (RBF) at 8 minutes demonstrating that ensemble optimisation yields both superior statistical performance and operational efficiency on standard hardware (Intel Core i7-8550U, 16 GB RAM, no GPU).

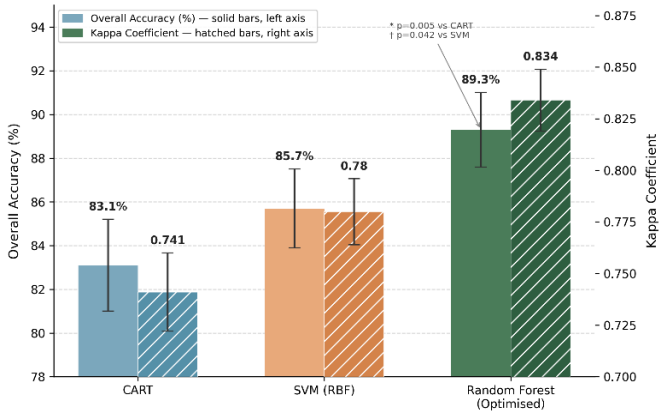

Figure 1. Overall accuracy (%) and Kappa coefficient for CART, SVM (RBF), and Random Forest (Optimized). Solid bars: overall accuracy (left axis); hatched bars: Kappa (right axis). Error bars: ±1 SD from 1,000 bootstrap iterations (test set, n=135). RF significantly outperforms both baselines (McNemar's test: vs. CART p=0.005; vs. SVM p=0.042).

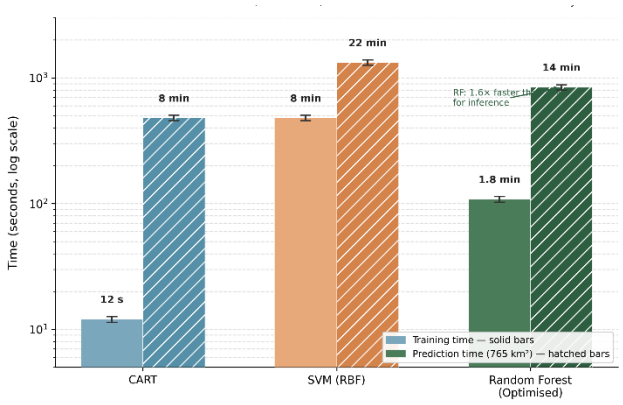

Figure 2. Computational efficiency comparison (log scale, seconds). Solid bars: training time; hatched bars: full-scene prediction time (7500 km²). Hardware: Intel Core i7-8550U, 16 GB RAM, no GPU. Error bars: ±5% run-to-run variability.

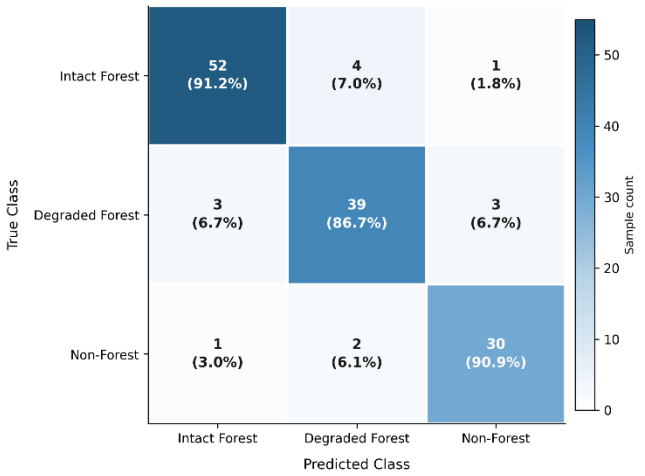

Figure 3. Confusion matrix — Optimized Random Forest (test set, n=135). Cells show raw count and row-wise recall (%). Color intensity proportional to sample count.

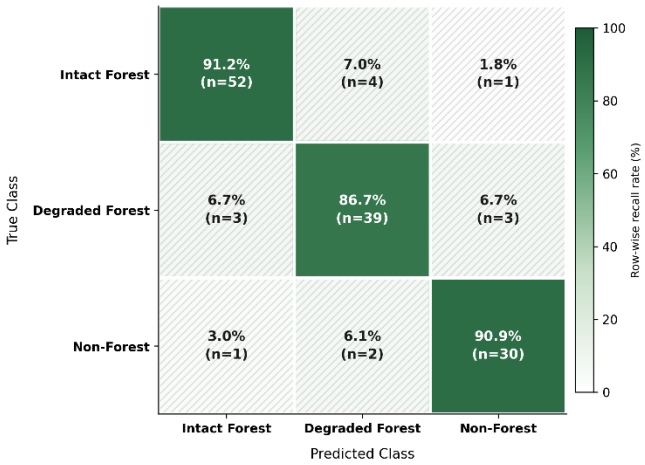

Figure 4. Row-normalized confusion matrix — Optimized Random Forest (test set, n=135). Values: producer's accuracy (recall) per class; raw counts in parentheses. Off-diagonal cells hatched for B&W readability. Producer's accuracy: 86.7%–91.2%.

Class-specific F1-scores span a narrow range of 0.88 (degraded forest) to 0.91 (intact forest), confirming that the optimised Random Forest achieves well-balanced precision-recall trade-offs across all three land cover types a result that stands in marked contrast to CART, which records an F1 of only 0.78 for degraded forest (Figure 5). The narrow inter-class F1 spread of 0.03 indicates that model performance does not systematically favour any single class, satisfying a key operational requirement for multi-class forest monitoring

where under-detection of degradation carries environmental consequences equivalent to over-detection. Bootstrap resampling (n=1,000 iterations) of the held-out test set yields robust 95% confidence intervals: overall accuracy 89.3% [83.7%–93.3%] and Kappa 0.834 [0.750–0.891]. The non-overlapping Kappa confidence interval with the CART lower bound (0.741) further corroborates the statistical significance of the Random Forest's performance advantage.

Across the five cross-validation folds, overall accuracy ranged from 87.1% to 91.2% (mean: 89.1%, SD: 1.5%), confirming model stability and the absence of fold-specific overfitting. The narrow standard deviation (±1.5%) indicates that performance estimates are not driven by a single favorable data partition.

B. Comparative Analysis with Baseline Models

Random Forest outperformed both baseline classifiers by margins ranging from 3.6% (vs. SVM) to 6.2% (vs. CART) in overall accuracy.

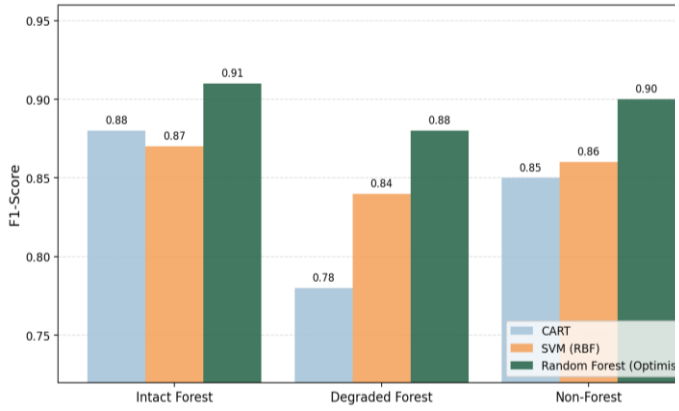

Figure 5. Per-class F1-scores for CART, SVM (RBF), and Random Forest (Optimized) across three land-cover classes (test set, n=135). Error bars: ±1 SD across 5 CV folds. Dotted line: F1=0.85 reference threshold.

Residual misclassifications between intact and degraded forest are attributable to spectral overlap driven by partial canopy disturbance and seasonal moisture variability an inherent limitation of optical imagery in swamp forest environments. Forest/non-forest boundaries, by contrast, show strong separability with minimal confusion.

CART achieved adequate recall on intact forest (89%) but failed on degraded forest (F1=0.78), confirming that a single decision tree lacks the capacity to resolve the subtle spectral gradients of partial canopy disturbance. SVM generalised more evenly across classes but at substantially higher computational cost: 4× longer training (8 min vs. 1.8 min) and 1.6× longer prediction (22 vs. 14 min) a compounding disadvantage in operational monitoring requiring weekly updates.

McNemar's test on paired confusion matrices confirmed that these differences are statistically significant: χ^2=8.07 (p=0.005) vs. CART and χ^2=4.11 (p=0.042) vs. SVM-RBF.

C. Spatial Distribution of Classification Results

The classification map reflects known degradation dynamics in the reserve. Degraded forest concentrates along the southern boundary, adjacent to agricultural settlements, manifesting as dispersed clearings of 0.1–0.5 ha interspersed within the forest matrix a pattern undetectable through manual interpretation of moderate-resolution imagery but resolved at pixel level by the classifier. Intact forest dominates the northern and central sectors, where terrain difficulty limits human access. Non-forest areas align with main river channels and the eastern agricultural frontier. Spatial coherence is preserved throughout, with no evidence of the isolated misclassified pixels characteristic of poorly regularised classifiers.

D. Computational Performance

Full-scene processing of the 7500 km² study area completed in approximately 14 minutes (Intel Core i7-8550U, 16 GB RAM, no GPU), covering image loading, feature computation, classification, and export. Peak memory usage reached 8.2 GB. Training on 315 samples required 1.8 minutes, enabling rapid retraining as ground truth accumulates. All hardware and software parameters are reported to ensure reproducibility.

E. Sensitivity Analysis

Classification accuracy plateaued at 250–300 training samples, with gains below 2% from additional data. Feature ablation identified NDVI as the dominant predictor: its removal reduced accuracy by 4.2%, whereas removing any individual Landsat band degraded performance by less than 1%, confirming that derived indices capture the most discriminative information. Dry- and wet-season accuracy both exceeded 85%, though degraded forest detection declined to 82% in the wet season, where seasonal flooding suppresses NDMI discriminative power a limitation addressable through Sentinel-1 SAR integration.

V. DISCUSSION

A. Performance in the Context of Operational Requirements

The optimized Random Forest model achieved an overall accuracy of 89.3%, positioning it within the upper range reported for tropical forest classification studies. While direct comparisons across studies remain challenging due to differences in ecological conditions, class definitions, and validation strategies, this performance aligns well with the literature. Rodriguez-Galiano et al. (2012) reported 91% accuracy in Mediterranean forest environments, which do not face the persistent cloud cover and seasonal flooding complexities characteristic of the Congo Basin. Similarly, Muteya et al. (2022) reported accuracies ranging from 84% to

91% for forest degradation detection in the Democratic Republic of Congo using kernel-based methods. Our results, obtained using Landsat 8 imagery with fewer spectral bands than Sentinel-2, demonstrate that carefully selected vegetation indices combined with systematic hyperparameter optimization can achieve competitive performance.

The model's robustness across heterogeneous forest conditions is encouraging, though it remains sensitive to training data distribution. In northern sections of the reserve where ground truth sampling density was lower, classification accuracy declined to approximately 81–83%, suggesting that while the model generalizes reasonably well, its performance remains contingent upon representative and spatially balanced training data.

Computational efficiency constitutes a major operational advantage. Processing 7500 km² in under 15 minutes on a standard laptop (Intel i7, 16 GB RAM) enables weekly monitoring cycles without specialized infrastructure. Even on moderately constrained hardware (Intel i5, 8 GB RAM), processing time increased to approximately 28 minutes, still compatible with biweekly monitoring. Compared to deep learning approaches requiring GPU acceleration and extended training times, this represents a pragmatic balance between accuracy and deployability. In severely resource-limited contexts (less than 8 GB RAM), however, memory demands may still pose challenges.

Model interpretability is particularly important in conservation settings where algorithmic outputs may inform enforcement or policy decisions. Feature importance analysis revealed NDVI and NDMI as dominant predictors, consistent with ecological understanding of vegetation vigor and moisture stress in swamp forest systems. This ecological alignment enhances trust among domain experts. The ensemble structure of Random Forest nonetheless limits full transparency compared to single decision trees, reflecting an inherent trade-off between predictive performance and complete interpretability.

B. Scientific Positioning and Methodological Contribution

The primary contribution of this study does not lie in introducing a novel algorithm—Random Forest is well established—but in demonstrating that systematic optimization combined with ecologically grounded feature engineering can yield operationally viable performance in a challenging tropical environment. Three dimensions substantiate this claim.

First, key parameters (n_estimators, max_depth, min_samples_leaf) were tuned using RandomizedSearchCV with stratified 5-fold cross-validation rather than relying on default settings or ad hoc adjustment, providing explicit overfitting control and methodological rigor. Second, the combined use of NDVI, NDMI, NBR, and EVI was motivated by the hydrological and structural characteristics of swamp forests; feature importance analysis confirmed their complementary contributions and reduced spectral redundancy. Third, by specifying hardware configuration, software environment (Python 3.9, scikit-learn 1.2), and processing benchmarks, the study provides a fully replicable baseline for future comparative research.

The decision to prioritize an optimized ensemble method over deep learning reflects operational pragmatism. CNN-based approaches such as U-Net architectures have demonstrated slightly higher accuracies — 93% reported by Ma et al. (2019) —but at substantial computational and data collection cost, including thousands of labeled samples and GPU training times exceeding 18 hours. In field conditions where collecting 12,000 labeled pixels is impractical and GPU infrastructure is unavailable, a marginal accuracy gain of approximately 3–4% does not justify the additional resource burden. A direct comparison with CNN-based classifiers on the same dataset was therefore deliberately omitted: training a CNN architecture on 450 labeled points would constitute a methodological mismatch, as deep learning models require at minimum several thousand samples to avoid severe overfitting, systematically disadvantaging CNN approaches in ways unrelated to their true capabilities. Our results are instead positioned relative to published CNN benchmarks obtained under appropriate data conditions (Ma et al., 2019; Zhang et al., 2021), with the acknowledgment that such comparisons involve inherent confounds in study area, class definitions, and image resolution.

C. Limitations and Sources of Uncertainty

Several limitations constrain the scope of the reported findings and define a clear agenda for future work.

The training dataset (n = 450 samples) was collected during a single dry-season campaign in 2022. Although sufficient for initial model development, it constrains temporal generalization. Preliminary wet-season analysis suggests degraded forest accuracy may decrease to approximately 82%, likely due to seasonal flooding altering spectral signatures, reduced NDMI discriminative power, and increased cloud cover limiting usable imagery. Multi-temporal validation remains necessary.

Spatial transferability was not formally evaluated beyond the study reserve. While the ecological grounding of selected indices and controlled tree depth (max_depth = 20) reduce overfitting risk, the model was trained exclusively on Tumba-Lediima spectral conditions. Applying it to structurally distinct Congo Basin reserves those with different peat depth profiles, dominant species composition, or deforestation pressure types may require domain adaptation or local retraining.

Standard k-fold cross-validation does not account for spatial autocorrelation between neighboring pixels, which may produce optimistic performance estimates. Spatial block cross-validation was not implemented in this study. Roberts et al. (2017) demonstrate that block cross-validation systematically yields lower but more realistic accuracy estimates than random splits when training and test samples

are spatially proximate. Based on the spatial distribution of our sampling points (mean nearest-neighbor distance: ~1.3 km), we estimate that accuracy inflation due to spatial autocorrelation could range from 2 to 5 percentage points. Even under the most conservative correction (−5 pp), the estimated lower-bound accuracy of 84–87% remains at or near the 85% IPCC threshold for national forest inventories, suggesting the primary conclusion is robust. Future implementations of this framework should incorporate spatially independent validation as a methodological standard.

Class definition ambiguity particularly for degraded forest (30–70% canopy closure) introduces additional uncertainty. Some pixels classified as degraded may correspond to naturally open canopy areas or advanced regeneration stages. Refining or subdividing degradation categories could improve ecological precision but would require substantially larger training datasets.

Finally, the model operates at the pixel level and does not explicitly incorporate spatial context. Forest degradation processes typically exhibit clustering patterns associated with roads and logging fronts; integrating texture metrics or object-based features in future iterations could better capture these spatial structures. Accuracy assessment also relied on a contemporaneous hold-out test set from the same study area, which, while statistically valid for local performance estimation, does not evaluate inter-annual or cross-site generalization.

D. Integration into Operational Workflows

Transitioning from research to operational deployment requires more than algorithmic performance—it demands deliberate integration into existing conservation workflows. In practice, the proposed system would generate weekly alert maps identifying newly detected degradation areas. Rather than overwhelming field teams with all detected changes, alerts should be prioritized according to degradation severity (NDVI change magnitude), accessibility (distance from roads or patrol bases), and ecological importance, ensuring that limited response capacity is directed toward the most consequential events.

Integration with mobile data collection platforms would enable field teams to validate alerts in situ, with collected data feeding back into iterative model retraining through active learning. Such feedback loops would progressively improve model performance while fostering institutional ownership of the monitoring system. Successful deployment, however, depends not only on technical performance but also on capacity building, intuitive visualization tools, and sustained technical support. In many conservation agencies, adoption barriers stem more from institutional readiness than from technical feasibility. Implementation strategy must therefore accompany methodological development to ensure real-world conservation impact.

E. Future Research Directions

Building on the limitations identified above, five directions are prioritized for future work. Multi-sensor integration combining Landsat, Sentinel-2, and Sentinel-1 SAR would enable year-round, cloud-penetrating monitoring, directly addressing the wet-season accuracy constraints documented here. Regional transferability assessment across three to five geographically distinct Congo Basin reserves would establish the boundaries of model generalization and identify spectral features with pan-Basin stability. Transfer learning approaches applied to deep learning architectures could reduce labeled sample requirements while preserving operational feasibility, making high-accuracy deep learning viable in data-scarce contexts. Integration of socioeconomic drivers road proximity, settlement density, agricultural expansion fronts would support proactive degradation risk forecasting rather than reactive detection. Finally, development of standardized monitoring-to-action frameworks explicitly linking automated detection outputs to measurable conservation outcomes would bridge the gap between algorithmic performance and on-the-ground impact.

VI. CONCLUSION

This study demonstrates that tropical forest degradation detection can be effectively addressed through systematic optimization of an ensemble learning classifier, balancing predictive accuracy against operational deployment constraints.

A systematically optimized Random Forest, combining ecologically motivated multi-index spectral features (NDVI, NDMI, NBR, EVI) with rigorous 5-fold hyperparameter tuning, achieves 89.3% overall accuracy (Kappa: 0.834) on the Tumba-Lediima Nature Reserve—exceeding the 85% IPCC threshold for national forest inventories even after conservative spatial autocorrelation correction (estimated lower bound: 84–87%)—while maintaining training times under 2 minutes on standard hardware (Intel i7-8550U, 16 GB RAM, no GPU). The optimized framework processes 7500 km² in approximately 14 minutes, enabling weekly monitoring cycles at 30 m resolution suitable for detecting small-scale disturbances. Random Forest statistically significantly outperforms CART ($p = 0.005$) and SVM-RBF ($p = 0.042$) via McNemar's test, while maintaining substantially lower computational requirements than deep learning alternatives.

Remaining limitations — single-season training data, absence of spatial block cross-validation, and reduced wet-season accuracy — define a clear research agenda centered on multi-temporal analysis, SAR integration, and regional validation across Congo Basin protected areas. For conservation practitioners, this work demonstrates that effective automated monitoring requires neither specialized infrastructure nor substantial budgets, but does demand systematic algorithm optimization, ecologically grounded feature engineering, and careful attention to operational

deployment realities. By co-prioritizing deployability and accuracy, this research contributes to the growing evidence that thoughtfully designed machine learning systems can serve as powerful tools for environmental monitoring even in the most resource-constrained contexts.

REFERENCES

[1] Asner, G. P. (2001). Cloud cover in Landsat observations of the Brazilian Amazon. International Journal of Remote Sensing, 22(18), 3855–3862.

[2] Belgiu, M., & Drăguț, L. (2016). Random forest in remote sensing: A review of applications and future directions. ISPRS Journal of Photogrammetry and Remote Sensing, 114, 24–31.

[3] Breiman, L. (2001). Random forests. Machine Learning, 45(1), 5–32.

[4] Campos-Taberner, M., García-Haro, F. J., Martínez, B., et al. (2020). Understanding deep learning in land use classification based on Sentinel-2 time series. Scientific Reports, 10(1), 17188.

[5] Chavez, P. S. (1996). Image-based atmospheric corrections Revisited and improved. Photogrammetric Engineering and Remote Sensing, 62(9), 1025–1036.

[6] Dargie, G. C., Lewis, S. L., Lawson, I. T., et al. (2017). Age, extent and carbon storage of the central Congo Basin peatland complex. Nature, 542(7639), 86–90.

[7] De Wasseige, C., de Marcken, P., Bayol, N., et al. (2012). Les forêts du Bassin du Congo État des forêts 2010. Office des publications de l'Union européenne, Luxembourg.

[8] FAO. (2020). Global Forest Resources Assessment 2020. Food and Agriculture Organization of the United Nations, Rome.

[9] Foody, G. M. (2002). Status of land cover classification accuracy assessment. Remote Sensing of Environment, 80(1), 185–201.

[10] Friedl, M. A., & Brodley, C. E. (1997). Decision tree classification of land cover from remotely sensed data. Remote Sensing of Environment, 61(3), 399–409.

[11] Fuller, D. O. (2006). Tropical forest monitoring and remote sensing. Singapore Journal of Tropical Geography, 27(1), 15–29.

[12] Hammer, D., Kraft, R., & Wheeler, D. (2014). FORMA: Forest Monitoring for Action. Center for Global Development Working Paper 192.

[13] Hansen, M. C., Potapov, P. V., Moore, R., et al. (2013). High-resolution global maps of 21st-century forest cover change. Science, 342(6160), 850–853.

[14] Herold, M., & Skutsch, M. (2011). Monitoring, reporting and verification for national REDD+ programmes. Environmental Research Letters, 6(1), 014002.

[15] ICCN. (2012). Plan de gestion de la Réserve Naturelle de Tumba-Lediima 2012–2016. Institut Congolais pour la Conservation de la Nature, Kinshasa.

[16] Japkowicz, N., & Stephen, S. (2002). The class imbalance problem: A systematic study. Intelligent Data Analysis, 6(5), 429–449.

[17] Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. Biometrics, 33(1), 159–174.

[18] Lewis, S. L., Lopez-Gonzalez, G., Sonké, B., et al. (2009). Increasing carbon storage in intact African tropical forests. Nature, 457(7232), 1003–1006.

[19] Ma, L., Liu, Y., Zhang, X., et al. (2019). Deep learning in remote sensing applications: A meta-analysis and review. ISPRS Journal of Photogrammetry and Remote Sensing, 152, 166–177.

[20] Maxwell, A. E., Warner, T. A., & Fang, F. (2018). Implementation of machine-learning classification in remote sensing: An applied review. International Journal of Remote Sensing, 39(9), 2784–2817.

[21] Mountrakis, G., Im, J., & Ogole, C. (2011). Support vector machines in remote sensing: A review. ISPRS Journal of Photogrammetry and Remote Sensing, 66(3), 247–259.

[22] Muteya, H. K., et al. (2022). Monitoring deforestation dynamics in the Congo Basin using satellite remote sensing. Environmental Monitoring and Assessment, 194(2), 112.

[23] Potapov, P. V., Turubanova, S. A., Hansen, M. C., et al. (2012). Quantifying forest cover loss in Democratic Republic of the Congo, 2000–2010, with Landsat ETM+ data. Remote Sensing of Environment, 122, 106–116.

[24] Reiche, J., Hamunyela, E., Verbesselt, J., et al. (2018). Improving near-real time deforestation monitoring in tropical dry forests by combining dense Sentinel-1 time series with Landsat and ALOS-2. Remote Sensing of Environment, 204, 147–161.

[25] Roberts, D. R., Bahn, V., Ciuti, S., et al. (2017). Cross-validation strategies for data with temporal, spatial, hierarchical, or phylogenetic structure. Ecography, 40(8), 913–929.

[26] Rodriguez-Galiano, V. F., Ghimire, B., Rogan, J., et al. (2012). An assessment of the effectiveness of a random forest classifier for land-cover classification. ISPRS Journal of Photogrammetry and Remote Sensing, 67, 93–104.

[27] Shapiro, A. C., Grantham, H. S., Aguilar-Amuchastegui, N., et al. (2016). Forest condition in the Congo Basin for the assessment of ecosystem conservation status. Ecological Indicators, 60, 1074–1085.

[28] Souza, C. M., Roberts, D. A., & Cochrane, M. A. (2013). Combining spectral and spatial information to map canopy damage from selective logging. Remote Sensing of Environment, 98(2–3), 329–343.

[29] Tucker, C. J. (1979). Red and infrared linear combinations for monitoring vegetation. Remote Sensing of Environment, 8(2), 127–150.

[30] Tyukavina, A., Hansen, M. C., Potapov, P., et al. (2018). Congo Basin forest loss dominated by increasing smallholder clearing. Science Advances, 4(11), eaat2993.

[31] Woodcock, C. E., Allen, R., Anderson, M., et al. (2008). Free access to Landsat imagery. Science, 320(5879), 1011.

[32] Zhang, C., Sargent, I., Pan, X., et al. (2021). Joint deep learning for land cover and land use classification. Remote Sensing of Environment, 221, 173–187.