REVISTA FLORESTA
Engenharia Florestal

ISSN 1982-4688
Edição 56
Universidade Federal do Paraná
Setor de Ciências Agrárias

ASSESSMENT OF FORESTRY AND LAND USE CARBON PROJECTS' EFFECTIVENESS IN THE VOLUNTARY MARKET

Tajudeen Okekunle Amusa[*1, 2], Timofeiczyk Junior Romano[2], Alexandre Behling[2]

[1]*Department of Forest Resources Management - University of Ilorin, Ilorin, Kwara State, Nigeria - amusa.to@unilorin.edu.ng; amusatajudeen@ufpr.br
[2]Department of Forest Engineering - Federal University of Paraná (UFPR), Curitiba, PR, Brazil - romano@ufpr.br; romano.timo@gmail.com alexandre.behling@gmail.com; alexandre.behling@yahoo.com.br

Received for publication: 21/09/2025 – Accepted for publication: 23/07/2026

Resumo

Avaliação da Efetividade dos Projetos de Carbono Florestal e de Uso da Terra no Mercado Voluntário. Projetos de carbono em Florestas e Uso da Terra (FUT) são importantes soluções baseadas na natureza para enfrentar as mudanças climáticas no Mercado Voluntário de Carbono (MVC). Entretanto, avaliações empíricas em larga escala sobre sua efetividade ainda são limitadas. Este estudo avaliou a efetividade de 732 projetos FLU certificados pelos principais padrões do mercado voluntário: Verra (VCS), Gold Standard, American Carbon Registry (ACR) e Climate Action Reserve (CAR). Os dados foram obtidos do banco harmonizado Berkeley Carbon Registry. Após o tratamento dos dados, aplicaram-se estatística descritiva, análise de agrupamento e regressão beta para identificar fatores associados à efetividade. Os resultados revelaram ampla variação entre registros, tipos de projeto e regiões. A análise de agrupamento identificou três tipologias de efetividade, sendo 94,4% dos projetos classificados como de escala média e moderadamente efetivos. A regressão beta (pseudo $R^2 = 0,16$) mostrou que a idade do projeto ($\beta = +0,48$; $p < 0,001$) e os créditos aposentados ($\beta = +0,59$; $p < 0,001$) aumentaram significativamente a efetividade, enquanto créditos emitidos ($\beta = -0,57$; $p < 0,001$), Gold Standard ($\beta = -1,31$; $p = 0,004$) e projetos REDD+ ($\beta = -0,44$; $p = 0,036$) a reduziram. O modelo estendido (pseudo $R^2 = 0,18$) revelou interações positivas da idade para Manejo Florestal Melhorado e Conversão Evitada de Pastagens. A efetividade prevista foi maior para A/R (0,53) e menor para REDD+ (0,43). Conclui-se que os projetos FLU apresentam efetividade moderada, sendo a idade do projeto e o volume de créditos aposentados os principais determinantes.

Palavras-chave: compensação de carbono; índice de aposentadoria; registro; tipologia de projeto; integridade de mercado

Abstract

Forestry and Land Use (FLU) carbon projects are important nature-based solutions to climate change within the Voluntary Carbon Market (VCM). However, large-scale empirical assessments of their real-world effectiveness remain limited. This study assessed the effectiveness of 732 FLU projects certified under four major voluntary carbon market standards: Verra (Verified Carbon Standard, VCS), Gold Standard, American Carbon Registry (ACR), and Climate Action Reserve (CAR). Project-level data were sourced from the harmonized Berkeley Carbon Registry Database. After data transformation, descriptive statistics, cluster analysis, and beta regression modeling were applied to identify structural and contextual drivers of effectiveness. Results showed wide variability in effectiveness by registry, project type, and region. Cluster analysis identified three effectiveness typologies, with the majority (94.4%) categorized as average-scale and moderately effective. The beta regression model (pseudo $R^2 = 0.16$) indicated that project age ($\beta = +0.48$, $p < 0.001$) and credits retired ($\beta = +0.59$, $p < 0.001$) significantly enhanced effectiveness, while credits issued ($\beta = -0.57$, $p < 0.001$), Gold Standard ($\beta = -1.31$, $p = 0.004$), and REDD+ projects ($\beta = -0.44$, $p = 0.036$) reduced it. An extended model (pseudo $R^2 = 0.18$) revealed positive age interactions for Improved Forest Management ($\beta = +0.43$) and Avoided Grassland Conversion ($\beta = +0.90$). Predicted effectiveness was highest for A/R (0.53) and lowest for REDD+ (0.43). These findings indicate that FLU projects operating within the VCM generally exhibit moderate levels of effectiveness, with project age and the volume of carbon credits retired being the most significant determinants.

Keywords: carbon offset; retirement ratio; registry; project typology; market integrity

INTRODUCTION

Forestry and land use (FLU) carbon projects occupy a central role in contemporary climate mitigation strategies. They have been promoted as cost-effective and nature-based solutions while simultaneously delivering co-benefits for biodiversity conservation, rural livelihoods, and sustainable development (GRISCOM *et al.*, 2017). Their implementation is increasingly mediated through the voluntary carbon market (VCM), which enables private actors to finance emissions reductions and removals by purchasing credits from FLU interventions. With a market value exceeding USD 2 billion (ECOSYSTEM MARKETPLACE, 2023), the VCM has become a dynamic platform for climate action, often positioned as a complementary pathway to state-led initiatives in tackling the scourge of climate change. It plays a particularly important role in channelling climate finance toward forest-rich regions of the Global South, where a number of FLU projects are located.

However, in recent years, concerns about the integrity of carbon offsets have intensified, with critics pointing to inflated baselines, weak monitoring, and limited community benefits (WEST *et al*., 2023). Early critiques, such as those by Michaelowa *et al*. (2019), revealed how flawed baseline assumptions and inadequate monitoring compromised the credibility of Clean Development Mechanism (CDM) projects, arguing that many initiatives failed to demonstrate true additionality. More recent assessments by Haya *et al*. (2020) and West *et al*. (2023) also show systematic over-crediting in REDD+ programmes, driven by inflated baselines and optimistic projections of forest loss. Their research revealed that many REDD+ credits did not correspond to actual emissions reductions, and that safeguards for forest communities were often weak or absent. These criticisms suggest that without robust institutional safeguards and context-sensitive project design, FLU interventions and others may fail to deliver durable climate benefits or equitable development outcomes.

With emerging global governance frameworks such as the Integrity Council for the Voluntary Carbon Market (ICVCM), robust empirical assessments of offset effectiveness are essential to inform reforms and build trust in market mechanisms. Despite the growing volume of research, there remains a limited body of rigorous, large-scale empirical studies that systematically analyse the effectiveness of FLU projects in measurable terms, particularly in the VCM. Much of the existing evidence is project-level, qualitative, or normative, offering insufficient insight into the structural and operational factors that influence project performance (GOLDSTEIN *et al*., 2021). This lack of robust empirical knowledge hampers both policy-making and market oversight, and undermines public trust in carbon trading schemes. The evidentiary gap poses serious implications for climate governance, carbon market design, and future investment in nature-based solutions (GRISCOM *et al*., 2017). Without robust empirical understanding of the factors that influence FLU project effectiveness, efforts to scale up forest carbon solutions risk being guided by assumptions rather than evidence (Schneider *et al*., 2019). Moreover, the absence of transparent and standardised performance benchmarks weakens accountability, impairs decision-making, and exposes carbon markets to reputational risks (HAMRICK AND GALLANT, 2018). There is thus a pressing need for analytical studies that evaluate effectiveness based on objective metrics.

In light of the foregoing, this study assessed effectiveness of FLU projects, measured as the ratio of carbon credits retired to credits issued. Specifically, the study sought to: (1) examine how institutional affiliation influences FLU project performance; (2) analyse regional variations in effectiveness outcomes across different geographical contexts; (3) investigate typological differences in performance based on project type; and (4) discuss the implications of these variations for climate finance, governance, and development outcomes, with a view to contributing to a more grounded understanding of what drives performance in forestry-based climate solutions. This focus responds to documented gaps in the literature (GOLDSTEIN *et al.*, 2021; HAYA et al., 2020), where existing studies are often project-specific, qualitative, or normative, with limited large-scale empirical analysis of effectiveness within the voluntary carbon market.

MATERIALS AND METHODS

Data collection

Using the Berkeley Voluntary Registry Offsets Database compiled by Haya et al. (2025), this study draws on a comprehensive project-level dataset to examine patterns of carbon credit issuance and retirement across voluntary market projects. The database, developed by the Berkeley Carbon Trading Project at the University of California, Berkeley, consolidates carbon offset records from four major VCM registries: Verra (VCS), Gold Standard, American Carbon Registry (ACR), and Climate Action Reserve (CAR). As of March 2025, the database contained 9,921 global carbon offset projects, of which 1,937 were identified as FLU interventions. Data were collected on registry, project type, carbon mechanism, geographic region, credits issued and retired, project age and effectiveness.

Data cleaning and pre-processing

Prior to statistical analysis and modelling, the dataset was cleaned and pre-processed to ensure analytical robustness and data reliability. Projects with missing or implausible records, including those with effectiveness ratios greater than 1 (i.e., credits retired exceeding credits issued), were excluded following standard data quality screening protocols for observational datasets. The final sample consisted of 732 valid FLU projects with complete records across key variables. These included: registry affiliation, project type (REDD+: Reducing Emissions from Deforestation and Forest Degradation are projects that aim to reduce greenhouse gas emissions by preventing deforestation and forest degradation, while promoting conservation and sustainable forest management, Afforestation/Reforestation (A/R): establishing forest cover on non-forested land (afforestation) or restoring forest on previously deforested areas (reforestation) to enhance carbon sequestration, Improved Forest Management (IFM): enhancement of carbon storage in existing forests through better silvicultural practices and reduced-impact

REVISTA FLORESTA
Engenharia Florestal
ISSN 1982-4688
Edição 56
Universidade Federal do Paraná
Setor de Ciências Agrárias

logging, Avoided Grassland Conversion: prevention of the transformation of native grasslands into croplands or other land uses, thereby preserving existing carbon stocks, Wetland Restoration: rehabilitation of degraded wetlands to restore their carbon sequestration capacity and ecological functions, Avoided Forest Conversion: deliberate actions aimed at preventing the planned or likely conversion of forests to non-forest uses, thus helping to maintain forest carbon stocks, Sustainable Grassland Management: applies to grazing control, reseeding, and conservation practices to maintain or enhance carbon in grassland ecosystems), carbon mechanism (reduction, removal, or mixed), geographic region, start year, credits issued, credits retired, remaining credits, and methodology used.

Each project was coded according to its institutional affiliation, intervention type, and regional location. Effectiveness was calculated as the ratio of credits retired to credits issued, serving as a proxy for performance (WEST *et al.*, 2023). Categorical variables were dummy-coded, and continuous variables such as credits and project age were log-transformed to address positive skewness and standardised using Z-scores for subsequent analysis (HAIR *et al.*, 2019).

Data analysis

A multi-tiered analytical strategy was employed to investigate the data collected. The analytical approach was structured around four sequential stages, each corresponding to a specific study objective.

In the first stage, descriptive statistics were used to summarize central tendencies and distributions across key variables. Measures included mean, median, standard deviation, trend (temporal) and skewness for continuous variables, and frequency distributions for categorical attributes (HAIR *et al.*, 2019). Cross-tabulations were used to explore patterns in effectiveness across registries, project types, and regions (AGRESTI, 2002). Effectiveness (E) was operationalised as:

$$E = \frac{\text{Credit retired}}{\text{Credit issued}} \text{.....................eqn.1 (WEST } et\ al.\text{, 2023)}$$

This ratio served as the primary performance proxy, consistent with approaches used in voluntary carbon market assessments (WEST *et al.*, 2023). It was normalized between 0 and 1.

In the second stage, inferential techniques were applied to test for statistically significant differences in effectiveness. One-way ANOVA was used to compare mean effectiveness across categorical groups, followed by Games-Howell post-hoc tests, which are appropriate when homogeneity of variance cannot be assumed. Additionally, Chi-square tests of independence were conducted to assess associations between institutional and regional variables and effectiveness categories (AGRESTI, 2002).

In the third stage, the K-means clustering was applied to uncover latent structures within the dataset, using standardized numerical features: project age, log-transformed credits issued and retired, and effectiveness ratio. The algorithm minimises the within-cluster sum of squares (WCSS):

$$WCSS = \sum_{K=1}^{K} \sum_{x_i \epsilon C_K} ||x_i - \mu_K||^2 \text{.....................eqn.2 (JAMES } et\ al.\text{, 2021)}$$

where x_i represents individual data points, μ_k is the centroid of cluster k, and K is the number of clusters.

The optimal number of clusters was determined using the Elbow Method, which identifies the point of diminishing returns in WCSS reduction, and the Silhouette Coefficient (JAMES *et al.*, 2021):

$$s(i) = \frac{b(i)-a(i)}{\max(a(i),b(i))} \text{.....................eqn.3}$$

where $a(i)$ is the mean intra-cluster distance and $b(i)$ is the mean nearest-cluster distance for observation i. Values approaching 1 indicate well-separated clusters.

The 3-cluster solution successfully grouped the 732 FLU projects into three distinct clusters with A one-way ANOVA confirming that credit-related variables (issued and retired) are the strongest discriminators among clusters.

In the fourth stage, beta regression modelling was employed to identify predictors of project effectiveness and quantify the strength and direction of relationships given that the dependent variable (effectiveness ratio) is continuous and bounded between 0 and 1 exclusive (KOSMIDIS AND ZEILEIS, 2024). The beta regression model is specified as:

$$g(\mu_i) = x_i{}^T\beta \text{ ....eqn.4}$$

where $g(\cdot)$ is a link function (logit), μ_i is the expected effectiveness for project i, x_i is the vector of predictors, and β is the vector of coefficients to be estimated.

The model included categorical predictors (registry type, project type, region) and continuous predictors (credits issued, credits retired, project age). A backward elimination process based on the likelihood ratio test (LRT) was used to refine the model:

$$LRT = -2\, In \left(\frac{L_{reduced}}{L_{full}}\right) \text{.....................eqn.5 (KOSMIDIS AND ZEILEIS, 2024)}$$

where $L_{reduced}$ and L_{full} are the likelihoods of the restricted and full models, respectively.

Model performance was evaluated using log-likelihood, Akaike Information Criterion (AIC; AKAIKE, 1974), pseudo R-squared values, and residual diagnostics.

An extended version of the model introduced an interaction term between project type and project age, revealing differential maturity-performance dynamics. Estimated marginal means (EMMs) were computed to assess predicted effectiveness across project types and regions (LENTH, 2016).

Software

All data analyses were conducted using Microsoft Excel (MICROSOFT CORPORATION, 2024), IBM SPSS Statistics Version 29 (IBM CORP., 2023), and R Version 4.3 (R CORE TEAM, 2024). Beta regression modelling was implemented using the *betareg* package (KOSMIDIS & ZEILEIS, 2024).

RESULTS

Descriptive overview of FLU projects in VCM

The forestry and land use (FLU) portfolio within the voluntary carbon market (VCM) exhibits considerable diversity in institutional affiliation, project typology, carbon accounting mechanisms, geographic distribution, and performance metrics. The analysis reveals that the Climate Action Reserve (CAR) hosts the largest share of FLU projects (41.94%), followed by the Verified Carbon Standard (VCS) at 34.29%, American Carbon Registry (ACR) at 21.72%, and Gold Standard (GS) with a modest 2.05%. Project typology is heavily skewed toward Improved Forest Management (IFM), which accounts for 63.39% of all entries. Afforestation/Reforestation (A/R) projects represent 17.21%, while REDD+ initiatives comprise 13.11%. Other categories, including Avoided Grassland Conversion, Avoided Forest Conversion, Wetland Restoration, and Sustainable Grassland Management, collectively account for less than 6.00% of the total. In terms of carbon mechanisms, mixed strategies dominate, with 65.03% of projects integrating both emission reduction and carbon removal components. Reduction-only projects make up 17.76%, while impermanent removal mechanisms, typically associated with A/R, constitute 17.21%. Geographically, the FLU project landscape is concentrated in the Global North and Latin America. North America leads with 41.26% of projects, closely followed by Latin America and the Caribbean at 39.75%. Sub-Saharan Africa, Eastern Asia, and South-Eastern Asia are underrepresented, hosting 7.38%, 6.56%, and 2.05% of projects respectively. Other regions, including Southern Asia, Oceania, and Europe, account for less than 3.00% combined (Table 1).

Table 1. Attributes of forestry and land use projects in VCM registries.
Tabela 1. Atributos de projetos florestais e de uso da terra em registros de VCM.

Attribute	Category	Frequency	Percentage (%)
Registry	CAR	307	41.94
	VCS	251	34.29
	ACR	159	21.72
	GOLD	15	2.05
Project Type	Improved Forest Management	464	63.39
	Afforestation/Reforestation	126	17.21
	REDD+	96	13.11
	Avoided Grassland Conversion	21	2.87
	Avoided Forest Conversion	13	1.78
	Wetland Restoration	7	0.96
	Sustainable Grassland Management	5	0.68
Mechanism	Mixed	476	65.03
	Reduction only	130	17.76
	Removal only	126	17.21
Region	North America	302	41.26
	Latin America and the Caribbean	291	39.75
	Sub-Saharan Africa	54	7.38
	Eastern Asia	48	6.56
	South-Eastern Asia	15	2.05
	Southern Asia	12	1.64
	Oceania	6	0.82
	Europe	2	0.27
	Eastern Europe	2	0.27

REVISTA FLORESTA Engenharia Florestal
ISSN 1982-4688
Edição 56
Universidade Federal do Paraná
Setor de Ciências Agrárias

Temporal analysis shows that project start years range from 1996 to 2023, with a mean vintage year of 2016.60. The age distribution is moderately right-skewed, with a mean of 8.42 years and a standard deviation of 4.48, indicating a predominance of relatively recent projects alongside a small cohort of older, legacy initiatives. Credit issuance and retirement patterns reveal significant variability (Figure 1). Issued credits range from 72 to nearly 40 million, with a mean of 1.13 million and a highly skewed distribution (skewness = 7.14; kurtosis = 58.96), driven by a few megaprojects. Retired credits follow a similar pattern, ranging from 0 to nearly 30 million, with a mean of 707,367 and comparable skewness and kurtosis values. These distributions underscore the heterogeneity in project scale and market uptake.

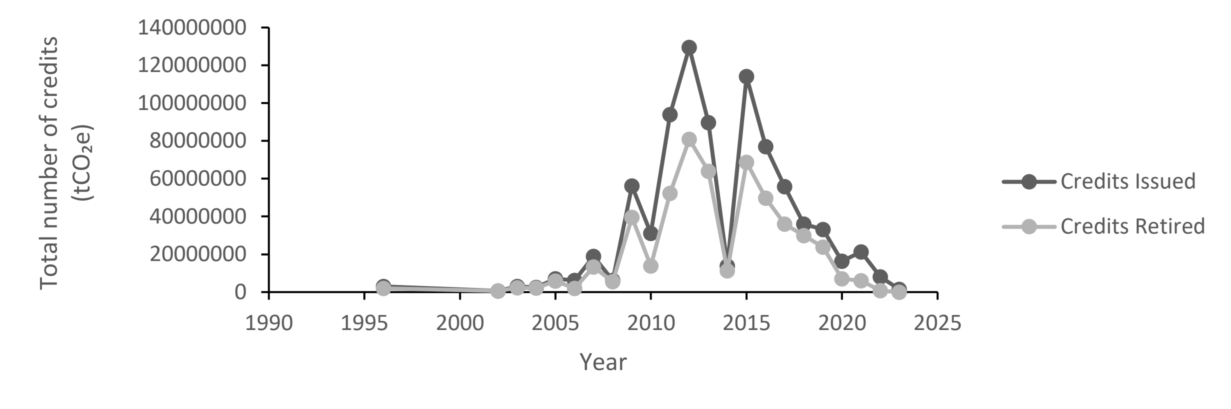

Figure 1. Trend in credits issued and retired for projects under forestry and land use.
Figura 1. Tendência de créditos emitidos e retirados para projetos de silvicultura e uso da terra.

Institutional Affiliation, Type and Effectiveness of FLU Projects

The Effectiveness Ratio for all 732 valid projects assessed, show a mean of 0.57, with a standard deviation of 0.39, near-symmetric distribution and a platykurtic profile (skewness = -0.39; kurtosis = -1.45). Projects registered under VCS recorded the highest mean effectiveness (0.61), followed closely by CAR (0.59; Table 2). ACR projects reported a lower average effectiveness of 0.49, while GOLD projects were the least effective on average, with a mean score of 0.43. These differences were statistically significant, as confirmed by a one-way ANOVA ($F(3, 728) = 4.219$, $p = 0.006$). Post hoc analysis using the Games-Howell test revealed that VCS projects were significantly more effective than those under ACR (mean difference = 0.120, $p = 0.011$), while other pairwise comparisons did not yield statistically significant differences. Further analysis of effectiveness by project type revealed additional variation. Avoided Forest Conversion projects had the highest average effectiveness (0.79), although this category included only thirteen projects. Afforestation/Reforestation (A/R) and REDD+ projects also performed relatively well, with mean effectiveness scores of 0.63 and 0.61 respectively. Improved Forest Management (IFM), which constituted the majority of the sample, had a moderate average effectiveness of 0.56. Other project types, including Avoided Grassland Conversion (0.45), Wetland Restoration (0.43), and Sustainable Grassland Management (0.30), recorded lower effectiveness scores. A one-way ANOVA confirmed that differences in effectiveness across project types were statistically significant ($F(6, 725) = 2.325$, $p = 0.029$). However, post hoc comparisons did not reveal statistically significant pairwise differences at the 0.05 level, likely due to small sample sizes in several categories.

Table 2. Mean effectiveness of forestry and land use projects by registry and type.
Tabela 2. Eficácia média dos projetos florestais e de uso da terra por registro e tipo.

Attributes		Frequency	Mean Effectiveness	Std. Deviation
Registry	VCS	251	0.61	0.34
	CAR	307	0.59	0.41
	ACR	159	0.49	0.40
	GOLD	15	0.43	0.38
Project Type	Afforestation/Reforestation	126	0.63	0.35
	Avoided Forest Conversion	13	0.79	0.24
	Avoided Grassland Conversion	21	0.45	0.40
	Improved Forest Management	464	0.56	0.41
	REDD+	96	0.61	0.32
	Sustainable Grassland Management	5	0.30	0.25
	Wetland Restoration	7	0.43	0.32

REVISTA FLORESTA
Engenharia Florestal
ISSN 1982-4688
Edição 56
Universidade Federal do Paraná
Setor de Ciências Agrárias

Regional patterns of effectiveness in FLU projects

Effectiveness of FLU projects within the VCM also varies across geographic regions. While North America and Latin America and the Caribbean host the majority of projects (302 and 291 respectively), their average effectiveness scores are only moderate (0.58 and 0.57 respectively; Table 3). In contrast, Oceania and Sub-Saharan Africa, though less represented in the dataset, recorded the highest mean effectiveness scores at 0.79 and 0.70 respectively. Projects in Europe reported an average effectiveness of zero, although this was based on only two entries with no credit retirements, limiting interpretive value. Other regions such as Eastern Asia (0.55), South-Eastern Asia (0.60), and Southern Asia (0.41) showed varied performance, with effectiveness scores clustering around the global mean of 0.57. A one-way ANOVA test indicated that the overall differences in effectiveness across regions were not statistically significant at the conventional threshold ($F(8, 723) = 1.906$, $p = 0.056$), though the p-value suggests a near-significant trend that warrants further exploration. Chi-square tests of projects distribution across effectiveness categories (Low, Medium, High) varied significantly by region, with a statistically significant association (χ^2 (16) = 27.54, $p = 0.036$), and a modest effect size (Cramer's V = 0.137).

Table 3. Mean effectiveness of forestry and land use projects by geographic region**.**
Tabela 3. Eficácia média dos projetos de silvicultura e uso da terra por região geográfica.

Attribute		Frequency	Mean Effectiveness	Std. Deviation
Region	Eastern Asia	48	0.55	0.36
	Eastern Europe	2	0.43	0.61
	Europe	2	0.00	0.00
	Latin America and the Caribbean	291	0.56	0.40
	North America	302	0.58	0.39
	Oceania	6	0.79	0.31
	South-Eastern Asia	15	0.59	0.33
	Southern Asia	12	0.41	0.35
	Sub-Saharan Africa	54	0.70	0.32

Cluster distribution of project effectiveness

The K-means clustering procedure partitioned the 732 FLU projects into three distinct and interpretable groups (Table 4). The majority of projects (n = 691; 94.4%) were assigned to Cluster 1, characterised by near-average values across all input variables, with slightly negative Z-scores for Total Credits Issued (−0.18), Total Credits Retired (−0.18), Project Age (−0.06), and Effectiveness (−0.02). Cluster 2 comprised a smaller subset of 35 projects (4.80%) and was markedly distinct, exhibiting substantially higher standardised scores for Total Credits Issued (1.98) and Total Credits Retired (2.10), alongside elevated Project Age (1.09) and Effectiveness (0.39). Cluster 3 was the most sparsely populated group, comprising just 6 projects (0.80%), and was defined by extraordinarily high Z-scores for both Total Credits Issued (8.79) and Total Credits Retired (8.55), with elevated but not extreme values for Project Age (1.06) and Effectiveness (0.14). Inter-cluster differences across the four variables were further assessed using one-way ANOVA (see Methods). Credit-related variables comprising credits issued and credits retired, emerged as the strongest discriminators among clusters ($F > 1{,}900$), followed by Project Age ($F = 27.34$), while Effectiveness exhibited the weakest discriminatory power ($F = 0.059$).

Table 4. Cluster summary: membership and performance characteristics (Z-scores).
Tabela 4. Resumo do cluster: características de associação e desempenho (pontuações Z).

Cluster	Cases (n)	% of Total (n = 732)	Total Credits Issued	Total Credits Retired	Project Age	Effectiveness
1	691	94.4%	-0.177	-0.180	-0.064	-0.021
2	35	4.8%	1.983	2.095	1.087	0.388
3	6	0.8%	8.792	8.550	1.061	0.135

Cluster composition by registry reveals that Cluster 1 spans all four registries, with CAR (303 projects) and VCS (224) most represented. Cluster 2 is more evenly distributed, though VCS remains dominant. Cluster 3 is exclusively composed of VCS projects, underscoring the platform's association with large-scale, high-credit operations. Cluster composition by project type shows that Cluster 1 includes nearly all intervention types, with Improved Forest Management (n = 452) most prevalent. Cluster 2 is concentrated in REDD+ (18 projects) and IFM (12), suggesting transitional or performance-enhancing conservation efforts. Cluster 3 consists solely of REDD+ projects, likely representing jurisdictional or flagship initiatives with exceptional scale.

Geographic distribution across clusters demonstrates significant regional differentiation. Cluster 1, which comprises the majority of projects, exhibits broad geographic representation. It is especially concentrated in North America (288 projects) and Latin America and the Caribbean (275), with additional presence in Eastern Asia (48) and Sub-Saharan Africa (47). Smaller numbers are found in South-Eastern Asia (11), Southern Asia (12), and Oceania (6).

Cluster 2, representing maturing or transitional projects, also spans multiple regions but is more concentrated in Latin America and the Caribbean (14) and North America (14). It includes a few entries from South-Eastern Asia (2) and Sub-Saharan Africa (5). Cluster 3 is sharply concentrated in Latin America and the Caribbean (2), South-Eastern Asia (2), and Sub-Saharan Africa (2).

Model performance and predictors of project effectiveness

A beta regression model fitted using maximum likelihood estimation converged successfully, yielding a log-likelihood value of 717.90 and a pseudo R^2 of 0.16 (Table 5). The model accurately captured variation in effectiveness without substantial systematic bias. There were no evidence of heteroscedasticity or model misspecification.

Table 5. Model fit statistics of predictors effectiveness of forestry and land use carbon projects.
Tabela 5. Estatísticas de ajuste do modelo de eficácia dos preditores de projetos de carbono florestal e de uso da terra.

Metric	Value
Log-likelihood	717.9
Degrees of Freedom	22
Pseudo R-squared	0.156
AIC	-1391.831
Estimation Method	Maximum Likelihood
Iterations	33 (BFGS) + 2 (Fisher scoring)

The total number of credits issued shows a strong and statistically significant negative association with project effectiveness (estimate = -0.57, $p < 0.001$). In contrast, the total number of credits retired has a positive and equally significant relationship with effectiveness (estimate = +0.58, $p < 0.001$). Similarly, project age is positively associated with effectiveness (estimate = +0.48, $p < 0.001$). Projects registered under the Gold Standard registry show a significantly lower effectiveness compared to those listed with other registries (estimate = -1.31, $p = 0.004$), while no significant differences were observed for CAR or VCS relative to the baseline category (ACR). REDD+ projects also exhibit significantly lower effectiveness than other project types such as afforestation/reforestation or improved forest management (estimate = -0.44, $p = 0.036$). None of the regional coefficients were statistically significant at the 5% level.

The fitted mean model with a logit link is given as:

$$\mathrm{logit}(\mu i) = 0.51 + 0.18 \cdot \mathrm{RegistryCAR} - 1.31 \cdot \mathrm{RegistryGOLD} - 0.20 \cdot \mathrm{RegistryVCS} - 0.22 \cdot \mathrm{TypeAFC} - 0.46 \cdot \mathrm{TypeAGC} - 0.15 \cdot \mathrm{TypeIFM} - 0.44 \cdot \mathrm{TypeREDD+} - \ldots - 0.57 \cdot \mathrm{TotalCreditsIssued} + 0.58 \cdot \mathrm{TotalCreditsRetired} + 0.48 \cdot \mathrm{Project.Age}$$

An extended model incorporating an interaction between project type and project age achieved a better fit. The model yield a higher log-likelihood of 725.50, an improved Pseudo R^2 of 0.18, and a lower AIC of -1395.09. The other interactions such as Region × Age or Registry × Age were explored but led to model non-convergence and multicollinearity issues due to data sparsity. Significant interaction effects were observed for Improved Forest Management × Age (estimate = +0.43, $p = 0.005$) and Avoided Grassland Conversion × Age (estimate = +0.90, $p = 0.027$).

Furthermore, estimated marginal means also show marginal effects by Project Type. Afforestation/Reforestation (A/R) projects demonstrate the highest estimated effectiveness (0.53; CI: 0.41 – 0.64), followed by Avoided Grassland Conversion (0.46; CI: 0.30 – 0.63), and Improved Forest Management (0.46; CI: 0.38 – 0.54). REDD+ projects show the lowest effectiveness (0.43; CI: 0.29 – 0.57), while Wetland Restoration projects has an estimated effectiveness of 0.44 (CI: 0.19–0.69).

DISCUSSION

Overview of FLU projects in VCM registries

The distribution of projects across registries shows the prominence of CAR and VCS in administering Forestry and Land Use (FLU) projects. Their dominance likely stems from earlier market entry, methodological flexibility, and extensive international partnerships. Conversely, the marginal presence of the Gold Standard suggests possible constraints in terms of project eligibility, geographic scope, or methodological specialization. These results align with trends identified in the literature (KOLLMUSS *et al*., 2015; HAMRICK AND GALLANT, 2018), where CAR and VCS consistently emerge as the leading platforms in VCM. Similarly, the predominance of Improved Forest Management (IFM) projects likely reflects preference for interventions perceived as manageable and scalable within existing silvicultural practices. By contrast, the underrepresentation of wetland and grassland initiatives, as suggested by GRISCOM *et al*. (2017) may be attributed to limited methodologies, technical complexity in carbon accounting, and the emerging nature of these approaches. The predominance of mixed-mechanism projects indicates an evolving sophistication in project design. This trend is indicative of an integrated landscape approach (SEYMOUR AND HARRIS, 2019), which has been recommended for maximizing both climate mitigation and biodiversity co-benefits. Meanwhile, the geographical skew in project distribution favouring North America and Latin America and Caribbean tends to reflect a confluence of factors including carbon market maturity, institutional readiness, registry affiliations, and policy frameworks supportive of carbon project implementation. The underrepresentation of Africa and Asia, despite their ecological potential may indicates persistent barriers such as limited access to financing, technical expertise, and robust governance systems. The temporal trend in number of projects under FLU suggests that the vast majority of project developments occurred within the last decade. This reflects a growing global interest in nature-based solutions

(NbS) and voluntary market mechanisms for climate mitigation. The skewness reinforces the observed surge in project registration post-2015, likely coinciding with the momentum generated by the Paris Agreement and rising corporate climate pledges. This trend demonstrates the dynamic evolution of the FLU carbon sector (IPCC, 2022), transitioning from early pilot efforts to a more mature and diverse global portfolio, increasingly aligned with broader climate and sustainable development goals

Institutional affiliation, project typology, and effectiveness of FLU initiatives

The mean effectiveness score of FLU projects (0.57) suggests that, on average, just over half of the carbon credits issued by such projects are ultimately retired. While this level of effectiveness may seem acceptable within a market-driven system (WEST *et al*., 2023; TRENCHER *et al*., 2024), it raises concerns when evaluated against the urgency of climate mitigation goals. The relatively modest average effectiveness score, coupled with high variability across registries and project types also has important implications for the integrity and credibility of the voluntary carbon market. The findings underscore the urgent need for harmonized effectiveness benchmarks and more transparent reporting of retirement rates. Moreover, the variability in project outcomes may be a pointer to a lack of consistency in implementation and oversight, which, if unaddressed, could undermine public and investor trust in nature-based climate solutions. The higher performance observed among projects certified by VCS and CAR relative to ACR and GOLD shows the presence of differential institutional capacities, verification rigour, and market reach. Hence, registries that enforce conservative baselines, ensure regular third-party verification, and require buffer pools or permanence safeguards are more likely to support effective projects. Conversely, registries perceived as lenient or lacking enforcement mechanisms may attract projects that underperform or over-promise (MICHAELOWA *et al*., 2019). This finding reinforces the idea that not all carbon credits are created equal, even when they derive from similar project types or geographies. It also highlights the need for meta-regulatory oversight or harmonization.

Regional effectiveness patterns in FLU initiatives

Our study revealed notable regional disparities in average project effectiveness. Underrepresented regions such as Sub-Saharan Africa (0.70) and Oceania (0.79) demonstrated relatively high performance, despite often facing systemic barriers such as limited technical expertise, insecure land tenure, and restricted access to finance. Meanwhile, regions with high project volumes, including North America (0.58) and Latin America and the Caribbean (0.56), though performed averagely, still showed comparatively lower effectiveness. This finding challenges a common assumption in carbon finance discourse, that project performance necessarily mirrors market maturity. On the contrary, it suggests that performance potential exists across regions (DUCHELLE *et al*., 2018), but is contingent upon enabling structures, including supportive governance, tailored methodologies, and capacity-building efforts.

Cluster-based insights into effectiveness typologies

The results on cluster analysis further underscore the importance of normalizing expectations in voluntary markets. Our findings show that not all projects are, or need to be, high-performing mega-offset schemes. The majority contribute incremental and verifiable benefits. This reflects insights from Boyd *et al.* (2007), who argue that the credibility of the carbon market rests not only on elite projects but on the reliability of its mainstream offerings. Cluster 1 projects as defined in our findings, might benefit most from targeted support aimed at performance enhancement, such as improved monitoring tools, access to third-party verification, and long-term funding stability. Many may be underperforming not due to conceptual flaws but because of operational constraints that could be mitigated with focused intervention. Similarly, projects in Cluster 2, characterized by moderate performance, revealed that effectiveness in them may only emerge over time when institutional maturity and adaptive learning are allowed to take root. This cluster supports the argument made by Duchelle *et al.* (2018), who highlight the slow-burn nature of forest-based mitigation efforts, particularly in complex socio-ecological settings. Cluster 3, though comprising less than 1.00% of the sample, holds significant symbolic and market weight. These projects exhibit exceptionally high effectiveness, large credit issuance and retirement volumes, and are exclusively REDD+ interventions registered under VCS. Geographically, they are concentrated in tropical forest regions of Latin America, Sub-Saharan Africa, and Southeast Asia. The prominence of these “flagship” projects speaks to both their carbon magnitude and their role as market influencers. However, while their results are impressive, that may not be easily generalizable due to contextual and methodological sensitivities. Importantly, the cluster framework shifts the discussion away from binary notions of “good” versus “bad” projects and instead promotes a graduated model of impact, rooted in context, process, and continuous improvement.

Key determinants of FLU project effectiveness

The positive and significant relationship between project age and effectiveness observed in this study confirms a widely held, yet insufficiently tested proposition in carbon market literature (KREIBICH AND OBERGASSEL, 2016), which indicate that older projects tend to perform better. Several mechanisms may underlie this pattern. First, longer operational timelines allow projects to stabilize institutional arrangements, complete verification cycles, and resolve early-stage implementation challenges. Second, over time, projects are more likely to achieve greater stakeholder buy-in, adaptive learning, and improved monitoring protocols, all of which contribute to increased credit retirement. This insight resonates with the work of Boyd *et al.* (2007), who emphasized the importance of “institutional stickiness” in forest carbon initiatives, i.e., the cumulative benefits of institutional continuity, stakeholder familiarity, and operational reliability.

In contrast to project age, the number of credits issued was negatively associated with effectiveness, indicating that scale does not equate to performance. This disconnect likely reflects structural issues in the voluntary carbon market (VCM). As noted by West *et al.* (2023), high issuance volumes may sometimes reflect generous or inflated baseline assumptions, such as in large REDD+ projects where deforestation avoidance is calculated against hypothetical future scenarios. In such cases, credits may be issued in anticipation of reductions that are difficult to verify or sustain, leading to a low retirement ratio and, consequently, low effectiveness. Another plausible explanation is that mega-projects tend to be more vulnerable to leakage, governance failure, or social contestation, all of which can undermine actual mitigation delivery. The scale of a project may increase exposure to political risk, corruption, or conflicting land uses, which are factors well-documented in critiques of large-scale offsetting schemes (TRENCHER *et al*., 2024).

Our findings have shown that retirement is the ultimate expression of carbon impact. This aligns with the emphasis in the literature on distinguishing between supply and demand dynamics in carbon markets. While issuance reflects supply-side potential, retirement reveals market demand and trust (HAMRICK AND GALLANT, 2018). Thus, projects with high retirement rates may benefit from superior reputational capital, better documentation, or verified co-benefits, which increase buyer confidence and uptake. This reinforces calls by Grassi *et al.* (2017) for performance-based standards that go beyond issuance volume or design ambition.

One of the more revealing outcomes of our results was the significant negative association between REDD+ projects and effectiveness, even after adjusting for project age and other factors. This suggests that, despite REDD+ being the most politically prominent and widely promoted project type in international climate finance, its carbon delivery performance lags behind other interventions such as Afforestation/Reforestation (A/R) or Improved Forest Management (IFM). The fact that REDD+ underperforms even in a global sample of certified projects may be a pointer to structural issues with its implementation, such as the difficulty of controlling deforestation drivers, weak enforcement of protection measures, and social resistance in contested landscapes.

Meanwhile, the finding that geographic region and carbon mechanism (removal vs. reduction vs. mixed) did not emerge as significant predictors in the main effects model once project-level variables were controlled for, suggests that effectiveness is not determined by where or how a project operates, but by how well it is implemented and maintained. This challenges geographic determinism in climate finance, i.e., the assumption that projects in certain countries or continents are inherently less credible, and supports a performance-based, rather than location-based, approach to carbon valuation and market trust. It also implies that carbon integrity can be achieved in diverse settings, provided that project governance, monitoring, and delivery mechanisms are robust.

CONCLUSION

- Forestry and land use (FLU) projects in the voluntary carbon market (VCM) generally show moderate effectiveness. For institutional affiliation, registry type significantly shapes project effectiveness, with projects under VCS and CAR generally showing higher effectiveness compared to ACR and Gold Standard.
- From regional context perspective, geographic location does not independently predict project effectiveness; instead, outcomes are strongly influenced by the quality of project implementation and verification.
- Effectiveness also varies across project typology/intervention types. Afforestation/Reforestation (A/R) and Improved Forest Management (IFM) tend to perform better, while REDD+ projects show comparatively lower effectiveness, reflecting differences in implementation complexity and baseline uncertainty.
- The key determinants of effectiveness were found to be project age and credits retired. This highlights the importance of operational maturity and verified delivery over nominal credit issuance.
- In terms of policy and market implications, our results show that current market practices that prioritize credit issuance as a proxy for impact need to be reconsidered. Greater emphasis should be placed on retirement-based performance metrics, transparency, and verification to strengthen market integrity and credibility.

ACKNOWLEDGEMENTS

The study was supported by Tertiary Education Trust Fund (TETFund), Nigeria, through a research fellowship sponsorship to the corresponding author in a tripartite partnership between TETFund, the Forum for Agricultural Research in Africa (FARA), and the Brazilian Association for International Education (FAUBAI).

REFERENCES

AGRESTI, A. Categorical data analysis. 2. ed. Hoboken: John Wiley & Sons, 2002.

AKAIKE, H. A new look at the statistical model identification. IEEE Transactions on Automatic Control, New York, v. 19, n. 6, p. 716–723, Dec. 1974.

BOYD, E.; MAY, P.; CHANG, M.; VEIGA, F. C. Exploring socioeconomic impacts of forest-based mitigation projects: lessons from Brazil and Bolivia. Environmental Science & Policy, Amsterdam, v. 10, n. 5, p. 419–433, Aug. 2007.

DUCHELLE, A. E.; SIMONET, G.; SUNDERLIN, W. D.; WUNDER, S. What is REDD+ achieving on the ground? Current Opinion in Environmental Sustainability, Amsterdam, v. 32, p. 134–140, Jun. 2018.

ECOSYSTEM MARKETPLACE; FOREST TRENDS. State of the voluntary carbon markets 2023. Washington, D.C.: Forest Trends, Sep. 2023. Available at: https://www.ecosystemmarketplace.com/publications/. Accessed on: 17 Jun. 2025.

GOLDSTEIN, A.; TURNER, W. R.; SPAWN, S. A.; ANDERSON-TEIXEIRA, K. J.; COOK-PATTON, S. C.; FARGIONE, J.; HOLE, D. G. Protecting irrecoverable carbon in Earth's ecosystems. Nature Climate Change, London, v. 10, n. 4, p. 287–295, Apr. 2020.

GRASSI, G.; HOUSE, J.; DENTENER, F.; FEDERICI, S.; DEN ELZEN, M.; PENMAN, J. The key role of forests in meeting climate targets requires science for credible mitigation. Nature Climate Change, London, v. 7, n. 3, p. 220–226, Mar. 2017.

GRISCOM, B. W.; ADAMS, J.; ELLIS, P. W.; HOUGHTON, R. A.; LOMAX, G.; MITEVA, D. A.; FARGIONE, J. Natural climate solutions. Proceedings of the National Academy of Sciences, Washington, D.C., v. 114, n. 44, p. 11645–11650, Oct. 2017.

HAIR, J. F.; BLACK, W. C.; BABIN, B. J.; ANDERSON, R. E. Multivariate data analysis. 8. ed. Boston: Cengage Learning, 2019.

HAMRICK, K.; GALLANT, M. Voluntary carbon market insights: 2018 outlook and first quarter trends. Washington, D.C.: Forest Trends' Ecosystem Marketplace, Apr. 2018. Available at: https://www.ecosystemmarketplace.com/publications/. Accessed on: 17 Jun. 2025.

HAYA, B. K.; BERNARD, T.; ABAYO, A.; RONG, X.; SO, I. S.; ELIAS, M. Voluntary Registry Offsets Database v2025-04. Berkeley: University of California, Berkeley Carbon Trading Project, Apr. 2025. Available at: https://gspp.berkeley.edu/faculty-and-impact/centers/cepp/projects/berkeley-carbon-trading-project/offsets-database. Accessed on: 1 Apr. 2025.

HAYA, B.; CULLENWARD, D.; STRONG, A. L.; GRUBERT, E.; HEILMAYR, R.; SIVAS, D. A.; WARA, M. Managing uncertainty in carbon offsets: insights from California's standardized approach. Climate Policy, Abingdon, v. 20, n. 9, p. 1112–1126, Sep. 2020.

INTERGOVERNMENTAL PANEL ON CLIMATE CHANGE (IPCC). Climate change 2022: mitigation of climate change. Cambridge: Cambridge University Press, 2022.

JAMES, G.; WITTEN, D.; HASTIE, T.; TIBSHIRANI, R. An introduction to statistical learning: with applications in R. 2. ed. Cham: Springer, 2021.

KARSENTY, A.; VOGEL, A.; CASTELL, F. "Carbon rights", REDD+ and payments for environmental services. Environmental Science & Policy, Amsterdam, v. 35, p. 20–29, Jan. 2014.

KOLLMUSS, A.; ZINK, H.; POLYCARP, C. Making sense of the voluntary carbon market: a comparison of carbon offset standards. Berlin: WWF Germany, Mar. 2008. Available at: https://www.worldwildlife.org/publications/. Accessed on: 17 Jun. 2025.

KOSMIDIS, I.; ZEILEIS, A. Extended-support beta regression for [0, 1] responses. Journal of the Royal Statistical Society: Series C (Applied Statistics), London, v. 75, n. 1, p. 139–157, Jan. 2024.

KREIBICH, N.; HERMWILLE, L. Robust performance metrics for results-based climate finance: recommendations for the Green Climate Fund. Wuppertal: Wuppertal Institute, 2016. Available at: https://epub.wupperinst.org/files/6264/6264_Carbon_Markets.pdf. Accessed on: 17 Jun. 2025.

LENTH, R. V. Least-squares means: the R package lsmeans. Journal of Statistical Software, Innsbruck, v. 69, n. 1, p. 1–33, Jan. 2016.

MICHAELOWA, A.; HERMWILLE, L.; OBERGASSEL, W.; BUTZENGEIGER, S. Additionality revisited: guarding the integrity of market mechanisms under the Paris Agreement. Climate Policy, Abingdon, v. 19, n. 10, p. 1211–1224, Dec. 2019.

NASI, R.; WUNDER, S.; CAMPOS, J. J. Forest ecosystem services: can they pay our way out of deforestation? Bogor: Center for International Forestry Research (CIFOR), 2011.

SCHNEIDER, L.; KOLLMUSS, A.; LAZARUS, M. Addressing the risk of double counting emission reductions under the Paris Agreement. Climatic Change, Dordrecht, v. 155, n. 2, p. 377–391, Mar. 2019.

SEYMOUR, F.; HARRIS, N. L. Reducing tropical deforestation. Science, Washington, D.C., v. 365, n. 6455, p. 756–757, Aug. 2019.

TRENCHER, G.; NICK, S.; CARLSON, J.; JOHNSON, M. Demand for low-quality offsets by major companies undermines climate integrity of the voluntary carbon market. Nature Communications, London, v. 15, n. 1, art. 6863, Aug. 2024.

WEST, T. A. P.; WUNDER, S.; SILLS, E. O.; BÖRNER, J.; RIFAI, S. W.; NEIDERMEIER, A. N.; FREY, G. P.; KONTOLEON, A. Action needed to make carbon offsets from forest conservation work for climate change mitigation. Science, Washington, D.C., v. 381, n. 6660, p. 873–877, Sep. 2023.