# The Effect of Atmospheric Correction Approaches on Leaf Area Index Retrieval Accuracy in Heterogeneous Croplands

**Sosdito Mananze**
Laboratory of Geographic Information Systems and Remote Sensing, Department of Rural Sociology, School of Rural Development, Eduardo Mondlane University, Mozambique | Regional Centre of Excellence in Agri-Food Systems and Nutrition, Eduardo Mondlane University, Maputo, Mozambique | United Methodist University of Mozambique, Cambine, Inhambane, Mozambique
sosdito_mananze@umum.ac.mz (corresponding author)

**Sebastiao Vilanculos**
Laboratory of Geographic Information Systems and Remote Sensing, Department of Rural Sociology, School of Rural Development, Eduardo Mondlane University, Mozambique | Regional Centre of Excellence in Agri-Food Systems and Nutrition, Eduardo Mondlane University, Maputo, Mozambique
sebastiaovilanculos81@gmail.com

**Mario Cunha**
Faculty of Sciences, University of Porto, Portugal | INESCTEC, Centre for Robotics in Industry and Intelligent Systems, Portugal
mccunha@fc.up.pt

*Received: 13 November 2025 | Revised: 23 December 2025 | Accepted: 3 January 2026*

 *DOI: https://doi.org/10.48084/etasr.16209*

## ABSTRACT

**This study evaluates four atmospheric correction approaches, including Sentinel-2 Correction (Sen2Cor), Rayleigh correction, image correction iCOR, and Dark Object Subtraction (DOS) for Leaf Area Index (LAI) retrieval in heterogeneous croplands in southern Mozambique, using the Sentinel Application Platform (SNAP) biophysical processor with PROSAIL neural network inversion. Field LAI measurements from 270 sample points across multiple vegetation types were used as reference data. The results showed that iCOR delivered the best overall performance, with a Root Mean Square Error (RMSE) of 1.315, a Mean Absolute Error (MAE) of 1.063, and a Bias of -0.684. This was followed by DOS, with an RMSE of 1.423, Sen2Cor with an RMSE of 1.514, and Rayleigh correction with an RMSE of 1.567. All methods exhibited negative coefficient of determination ($R^2$) values from-1.223 to -2.155, indicating systematic LAI underestimation and limited predictive capability. Performance varied substantially across stratification levels. LAI-class analysis revealed that DOS performed the best for low LAI with an RMSE of 0.478, while iCOR excelled at medium LAI with an RMSE of 0.956, and high LAI with an RMSE of 1.929. Strong positive correlations between absolute error and field LAI (Pearson's correlation coefficient ($r$) = 0.733 to 0.895) indicated increasing retrieval challenges at higher canopy densities. Significant spatial variation was observed, with ESUDER Campus exhibiting lower errors (RMSE = 0.518-0.745) compared to Machengue (RMSE = 1.435-1.741). Vegetation-type specific analysis showed that Miombo Forest had the lowest errors (RMSE = 0.542 for DOS), while Mango orchards presented the greatest challenges (RMSE = 1.754 for iCOR). These findings demonstrate the importance of selecting appropriate atmospheric correction methods based on specific vegetation characteristics, canopy density, and local environmental conditions for operational LAI monitoring in heterogeneous agricultural landscapes.**

***Keywords-leaf area index; sentinel-2; atmospheric correction; heterogeneous agricultural landscapes***

## I. INTRODUCTION

LAI represents one of the most important biophysical parameters for understanding vegetation dynamics and ecosystem functioning [1]. As an Essential Climate Variable (ECV) recognized by the Global Terrestrial Observing System (GTOS), LAI applications span from vegetation growth monitoring to the modeling of water, carbon, and energy fluxes between land and atmosphere [2-4]. Several approaches have been proposed for the estimation of LAI based on remote sensing data. These approaches can be grouped into parametric and non-parametric algorithms, physically-based methods, and the inversion of radiative transfer models with look-up tables [5-7]. Applying these methods resulted in global LAI products of varying spatial resolutions, including LAI products from Moderate Resolution Imaging Spectroradiometer (MODIS) [4], Advanced Very-High-Resolution Radiometer (AVHRR) [8], Satellite Pour Observation de la Terre Vegetation (SPOT-VGT), Proba-V [3], and Sentinel-2 [9].

Atmospheric correction quality can significantly influence downstream biophysical parameter retrieval. Sensitivity analyses reported in [10] show that LAI estimation accuracy varies according to canopy cover type and spectral variability, particularly when radiative transfer model inversions are applied. These results emphasize the need for systematically evaluating atmospheric correction approaches across diverse vegetation conditions and canopy densities, a methodological gap that directly motivated the present investigation in heterogeneous tropical croplands.

Given their high spatial resolution (10 m), Sentinel-2 products are promising for more accurate and reliable agricultural monitoring as well as for application in precision farming management. They have been used for numerous agricultural applications, including cropland mapping [11] and crop growth dynamic mapping [12]. Nevertheless, studies aiming to validate Sentinel-2-derived LAI against in-situ LAI data report different levels of agreement. For example, authors in [5] found good performance over homogeneous canopies but poor performance in heterogeneous canopies in the US. Authors in [13] reported low agreement in large-scale and small-holder commercial cultivation areas of Maize and Sunflower in South Africa, while authors in [14] described less reliable estimates of LAI in large fields dominating areas in Poland. Authors in [9] found good correspondence in the main European Rice areas in Spain, Italy, and Greece. Authors in [15] reported underestimation of LAI over different annual crops in Canada, and authors in [6] reported low agreement in small-holder Maize fields in Mozambique. These discrepancies arise from multiple sources, including sensor characteristics, atmospheric correction procedures, retrieval algorithms, and significant differences in measurement scales and methodologies [11]. Authors in [13] revealed differences of LAI derived from Bottom-of-Atmosphere (BOA) and Top-of-Atmosphere (TOA) Sentinel-2 imagery, suggesting the effect of atmospheric correction. Accordingly, the present study investigates the effect of different approaches of atmospheric correction on the accuracy of Sentinel-2 LAI estimates through global vegetation biophysical variables retrieval algorithms.

## II. METHODS

### A. Study Area Description

The current study was conducted in the district of Vilankulo, southeast Mozambique. The climate in the area is dry tropical, typical of Mozambique's southern coastal region. Temperatures remain high year-round, with distinct seasons: a hot, humid season (November–March) and a cooler, dry season (April–October). Sandy soil is predominant in the district, with clay soils occurring along the riverine margins and seasonally inundated areas.

A total of 270 sampling points were systematically distributed across different vegetation types, including crops, fruit trees, and natural vegetation, maintaining an average distance of 30 m between points to ensure spatial representativeness. Field measurements were conducted in three distinct agricultural zones with 203 sampling points in the Machengue zone, 58 sampling points in the ESUDER Campus zone, and 9 sampling points in the Pambarra zone. Figure 1 presents the study area with three field collection sites.

### B. In-Situ LAI Collection

The in-situ LAI was measured using the LAI-2200C device (LI-COR Biosciences), which estimates leaf area per unit of ground surface based on the interception of optical radiation at five distinct zenith angles (7°, 23°, 38°, 53°, and 68°). To minimize interference from diffuse solar radiation, measurements were conducted during the evening twilight or within 2 h after sunrise, following standard protocols. Table I summarizes the descriptive statistics of LAI by vegetation class.

TABLE I. DESCRIPTIVE STATISTICS OF LAI VALUES FOR DIFFERENT CROPS

| Crops | Count | Min. | Max. | Std. deviation | Coeff. of variation (%) |
|---|---|---|---|---|---|
| Beans | 8 | 3.9 | 5.07 | 0.4 | 7.89 |
| Cabbage | 7 | 0.49 | 6.08 | 1.94 | 31.91 |
| Carrot | 3 | 2.34 | 4.17 | 0.94 | 22.54 |
| Grass | 5 | 3.36 | 4.36 | 0.45 | 10.32 |
| Maize | 80 | 0.89 | 4.83 | 1.11 | 22.98 |
| Mango tree | 26 | 3.38 | 5.33 | 0.56 | 10.51 |
| Mixed crops | 45 | 1.14 | 5.99 | 1.17 | 19.53 |
| Okra | 3 | 2.56 | 5.13 | 1.3 | 25.34 |
| Peanut | 4 | 1.64 | 3.42 | 0.78 | 22.81 |
| Sweet potato | 20 | 1.13 | 5.79 | 1.19 | 20.55 |
| Tomato | 10 | 0.82 | 6.11 | 1.78 | 29.13 |
| Various plants | 38 | 1.68 | 5.24 | 0.72 | 13.74 |
| Watermelon | 19 | 2.81 | 4.35 | 0.49 | 11.26 |

### C. Sentinel-2 Data Acquisition

A total of four Sentinel-2 Level-1C (TOA reflectance) images, covering areas of interest, were acquired from the Copernicus Open Access Hub. The images were selected based on cloud cover (<10%) and temporal alignment with the field data collection period. Table II presents the details of each acquired image.

TABLE II. DETAILS OF SENTINEL-2 MSI LEVEL-1C TOA IMAGES USED IN THE STUDY

| Scene ID | Date |
|---|---|
| S2B_MSIL1C_20230709T072619_N0509_R049_T36KYA | 09/07/2023 |
| S2A_MSIL1C_20240330T072611_N0510_R049_T36KYA | 30/03/2024 |
| S2B_MSIL1C_20240822T072619_N0511_R049_T36KYA | 22/08/2024 |
| S2A_MSIL1C_20240926T072651_N0511_R049_T36KYA | 26/09/2024 |

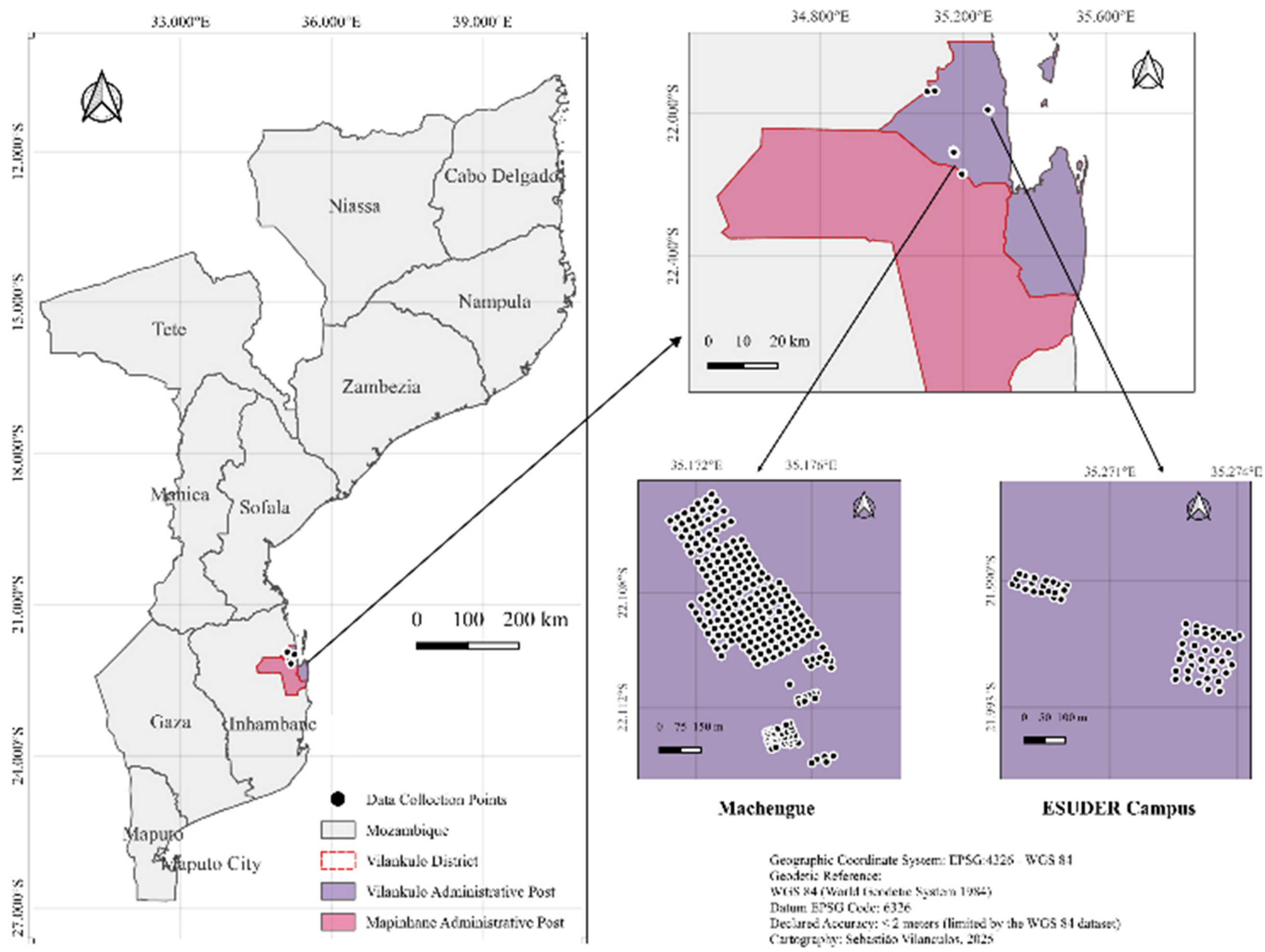


Fig. 1. Study area with three field collection sites.

### *D. Sentinel-2 Data Pre-Processing*

The pre-processing of the Sentinel-2 Level-1C (TOA reflectance) images involved the application of four atmospheric correction methods, including Sen2Cor, iCOR, Rayleigh correction, and DOS.

The Sen2Cor is an algorithm [16] developed by the European Space Agency (ESA), which performs the conversion to Level-2A (BOA) through the modeling of atmospheric parameters, such as water vapor and ozone, complemented by topographic adjustments based on the Shuttle Radar Topography Mission (SRTM) model. The iCOR algorithm, developed in [17] and implemented as a plugin for SNAP by the Vlaamse Instelling voor Technologisch Onderzoek (VITO), is particularly suitable for mixed scenarios (land/water), reducing spectral discontinuities between terrestrial and aquatic targets. iCOR combines MODTRAN 5-based atmospheric modeling, using pre-calculated reference look-up tables for parameters such as aerosol optical thickness and water vapor; automatic detection of water pixels (threshold in the B08 band) and clouds (adjustable thresholds in the coastal and VNIR bands); and correction of adjacency effects via SIMEC, essential for dynamic water bodies [18]. According to [16], iCOR stands out for its ability to adapt to different multispectral sensors and for minimizing systematic errors in coastal regions. Rayleigh correction is significant for processing data from Sentinel-2, removing atmospheric molecular scattering that distorts TOA reflectance, especially in blue bands. It is based on radiative transfer models and calculates BRR reflectance, considering parameters such as Rayleigh Optical Thickness (ROT) and acquisition angles. Although effective (reduces ~60% of the TOA signal in blue), it has limitations because it assumes a simplified atmosphere and Lambertian surface [19]. The DOS was implemented under the Semi-Automatic Classification Plugin [20] in the QGIS environment and performs quick corrections in case of computational limitations.

### *E. Retrieval of LAI from Sentinel-2*

Each atmospherically corrected image was used to estimate Sentinel-2 LAI. The estimation was performed employing biophysical processing tools within the SNAP environment. These tools utilize a neural network inversion technique trained on simulations from the PROSAIL radiative transfer model to derive LAI values from canopy-top reflectance. The estimated Sentinel-2 LAI was compared to the field LAI, aiming to investigate the effect of different atmospheric correction approaches on the performance of LAI retrieval from Sentinel-2 imagery.

*F. Normalization Procedure*

Field LAI data were normalized to account for the scale difference between field and satellite-derived values and enable a meaningful comparison. An NDVI-derived Fraction of Vegetation Coverage (FVC) was used to normalize the field LAI data using:

$$FVC = \frac{NDVI - NDVI_{soil}}{NDVI_{veg} - NDVI_{soil}} \quad (1)$$

where $NDVI_{soil}$ and $NDVI_{veg}$ represent bare soil and full vegetation coverage conditions of the study areas, respectively. This normalization accounts for site-specific soil background effects and vegetation coverage variations.

*G. Statistical Analysis*

Satellite-derived LAI estimates from each atmospheric correction method were validated against field-measured LAI using standard statistical metrics. Bias was calculated as the mean difference between satellite and field LAI, indicating systematic over- or underestimation. MAE and RMSE quantified the magnitude of deviations. $R^2$ assessed the proportion of variance in field LAI explained by satellite estimates, while Pearson's correlation coefficient ($r$) measured the linear association between the observed and predicted values.

To investigate the influence of vegetation characteristics on retrieval accuracy, performance metrics were calculated across three stratification levels. First, samples were classified into LAI terciles (Low: LAI ≤ 1.31, Medium: 1.31 < LAI ≤ 2.08, High: LAI > 2.08) to examine sensitivity to canopy density. Second, vegetation-type specific analysis was conducted for categories with adequate sample size ($n \geq 20$): maize ($n = 80$), mango tree ($n = 26$), Miombo forest ($n = 38$), Mixed crops ($n = 45$), and sweet potato ($n = 20$). Third, spatial variability was assessed by comparing performance between the two primary sampling locations, Machengue ($n = 211$) and ESUDER Campus ($n = 59$).

The relationship between retrieval accuracy and field LAI magnitude was quantified using Pearson's correlation between absolute error and observed field LAI. This analysis identified whether systematic error patterns varied with canopy density, as proposed in [10] for radiative transfer inversions in diverse vegetation conditions.

## III. RESULTS

*A. Overall Validation Performance*

Validation of satellite-derived LAI against field measurements revealed substantial differences in performance among the four atmospheric correction methods, as shown in Table III and Figure 2. iCOR demonstrated the best overall accuracy with an RMSE of 1.315, an MAE of 1.063, and a Bias of -0.684, indicating the lowest systematic underestimation. DOS ranked second with an RMSE of 1.423 and a Bias of -0.995. This was followed by Sen2Cor with an RMSE of 1.514 and a Bias of -0.934, and Rayleigh correction with an RMSE of 1.567 and a Bias of -1.185.

All four methods exhibited negative $R^2$ values, ranging from -1.223 (iCOR) to -2.155 (Rayleigh correction), indicating that those predictions performed worse than a horizontal line at the mean field LAI. Pearson correlation coefficients were near zero or slightly negative (-0.131 to -0.172), confirming the absence of meaningful linear relationships between field and satellite LAI across the full dataset.

TABLE III. GENERAL VALIDATION METRICS FOR LAI RETRIEVAL (N = 270)

| Method | Bias | MAE | RMSE | $R^2$ | $r$ |
|---|---|---|---|---|---|
| iCOR | -0.684 | 1.063 | 1.315 | -1.223 | -0.166 |
| DOS | -0.995 | 1.136 | 1.423 | -1.603 | -0.172 |
| Sen2Cor | -0.934 | 1.230 | 1.514 | -1.944 | -0.146 |
| Rayleigh correction | -1.185 | 1.276 | 1.567 | -2.155 | -0.131 |

*B. Sensitivity to LAI Magnitude*

Stratification by LAI terciles revealed pronounced sensitivity of retrieval accuracy to canopy density, as presented in Table IV. For low LAI (LAI ≤ 1.31, $n = 90$), DOS achieved the best performance with an RMSE of 0.478. This was followed by Rayleigh correction with an RMSE of 0.547, iCOR with an RMSE of 0.744, and Sen2Cor with an RMSE of 0.762. However, in the medium LAI range (LAI = 1.31-2.08, $n = 90$), iCOR became the most accurate method, achieving an RMSE of 0.956, with DOS ranked second obtaining an RMSE of 1.095, Sen2Cor third with an RMSE of 1.184, and Rayleigh correction performing the worst with an RMSE of 1.278. This ranking persisted for high LAI (LAI > 2.08, $n = 90$), where iCOR maintained superiority (RMSE = 1.929), followed by DOS (RMSE = 2.156), Sen2Cor (RMSE = 2.212), and Rayleigh correction (RMSE = 2.331).

TABLE IV. VALIDATION METRICS BY LAI CLASS ($n = 90$)

| LAI class | Method | Bias | MAE | RMSE | $R^2$ | $r$ |
|---|---|---|---|---|---|---|
| Low | Sen2Cor | 0.138 | 0.568 | 0.762 | -5.682 | -0.023 |
| | Rayleigh | -0.171 | 0.442 | 0.547 | -2.441 | -0.037 |
| | iCOR | 0.387 | 0.609 | 0.744 | -5.36 | -0.032 |
| | DOS | 0.025 | 0.394 | 0.478 | -1.624 | -0.015 |
| Medium | Sen2Cor | -0.953 | 1.085 | 1.184 | -27.371 | -0.157 |
| | Rayleigh | -1.186 | 1.188 | 1.278 | -32.053 | -0.152 |
| | iCOR | -0.714 | 0.846 | 0.956 | -17.507 | -0.155 |
| | DOS | -0.99 | 0.994 | 1.095 | -23.293 | -0.174 |
| High | Sen2Cor | -1.987 | 2.035 | 2.212 | -12.092 | -0.087 |
| | Rayleigh | -2.199 | 2.199 | 2.331 | -13.544 | -0.071 |
| | iCOR | -1.724 | 1.735 | 1.929 | -8.962 | -0.117 |
| | DOS | -2.021 | 2.021 | 2.156 | -11.441 | -0.129 |

A strong positive correlation was observed between the absolute error and field LAI magnitude, as presented in Table V, with correlation coefficients ranging from 0.733 (iCOR) to 0.895 (Rayleigh). This pattern indicates that all methods experienced increasing difficulty retrieving LAI accurately as canopy density increased, with errors systematically growing at higher LAI values.

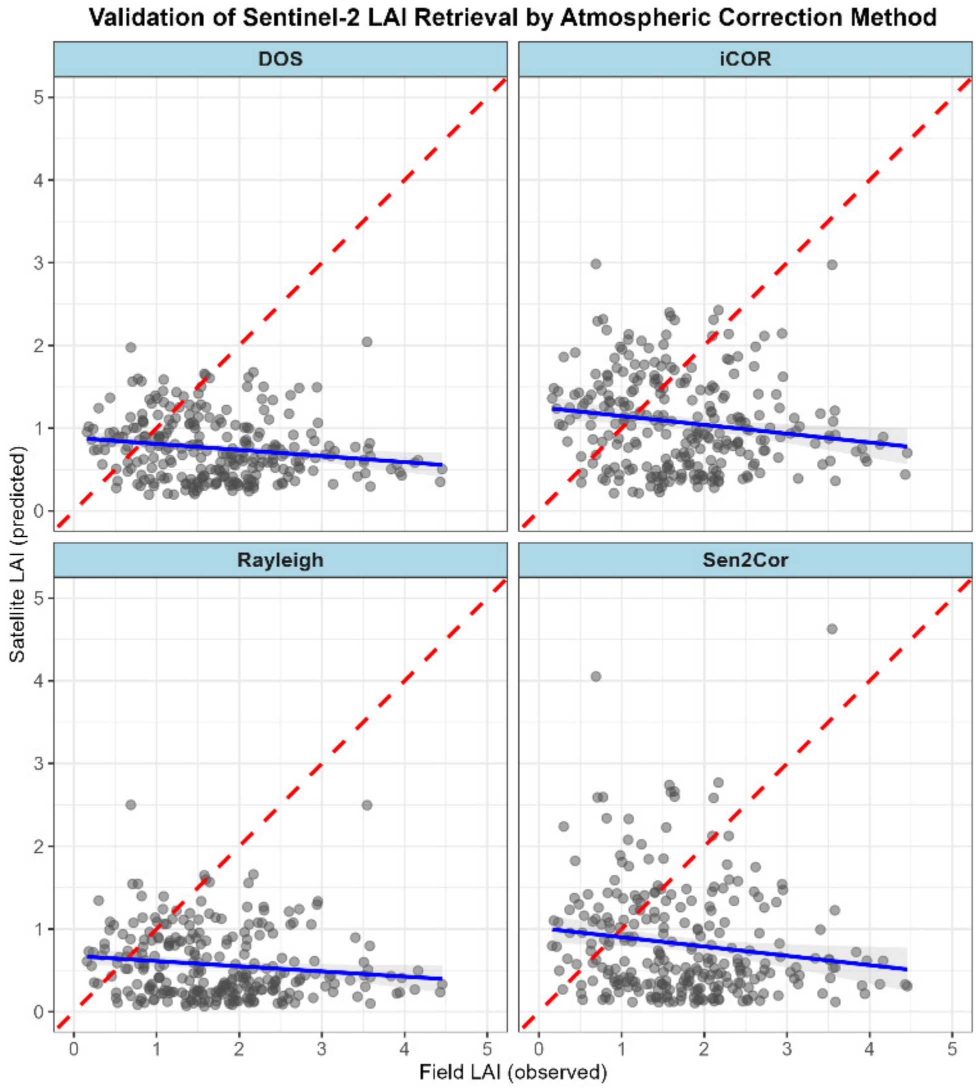


Fig. 2. Validation of satellite LAI retrieval by atmospheric correction methods.

TABLE V. CORRELATION BETWEEN ABSOLUTE ERROR AND FIELD LAI

| Method | *r* |
|---|---|
| iCOR | 0.733 |
| Sen2Cor | 0.799 |
| DOS | 0.890 |
| Rayleigh | 0.895 |

## C. *Variation by Location and Vegetation Type*

As illustrated in Table VI, spatial analysis revealed significant differences in retrieval performance between the two primary sampling locations. At the ESUDER Campus zone, all methods performed substantially better, with RMSE values ranging from 0.518 for DOS to 0.745 for iCOR. DOS achieved a positive $R^2$ of 0.194 at this location, indicating some predictive capability. In contrast, the Machengue zone exhibited significantly higher errors, with RMSE ranging from 1.435 for iCOR to 1.741 for Rayleigh correction, and all $R^2$ values remaining negative.

TABLE VI. VEGETATION-TYPE SPECIFIC ANOVA RESULTS

| Location | Method | Bias | MAE | RMSE | $R^2$ | *r* |
|---|---|---|---|---|---|---|
| ESUDER Campus zone | DOS | 0.003 | 0.414 | 0.518 | 0.194 | 0.442 |
| | Rayleigh correction | -0.282 | 0.482 | 0.627 | -0.181 | 0.333 |
| | Sen2Cor | 0.275 | 0.585 | 0.732 | -0.608 | 0.348 |
| | iCOR | 0.522 | 0.633 | 0.745 | -0.667 | 0.424 |
| Machengue zone | iCOR | -1.021 | 1.184 | 1.435 | -1.549 | -0.005 |
| | DOS | -1.274 | 1.339 | 1.587 | -2.117 | -0.005 |
| | Sen2Cor | -1.272 | 1.410 | 1.668 | -2.446 | -0.036 |
| | Rayleigh correction | -1.438 | 1.499 | 1.741 | -2.755 | -0.024 |

Vegetation-type specific analysis, depicted in Table VII, demonstrates additional variability in method performance. Miombo forest yielded the lowest errors across all vegetation types, with DOS achieving an RMSE of 0.542. Maize croplands also showed relatively good performance, with iCOR reaching an RMSE of 0.990. Conversely, Mango orchards presented the greatest retrieval challenges, with iCOR producing an RMSE of 1.754 despite being the best-performing method for this vegetation type. Mixed crop systems with an RMSE of 1.373 for iCOR and sweet potato fields with an RMSE of 1.344 for iCOR exhibited intermediate error levels. Across most vegetation types except the Miombo forest, iCOR delivered the best performance.

TABLE VII. VALIDATION METRICS BY VEGETATION TYPE ($n \geqslant 20$)

| Vegetation type | Method | Bias | MAE | RMSE | $R^2$ | r |
|---|---|---|---|---|---|---|
| Maize | iCOR | -0.232 | 0.802 | 0.990 | -0.457 | -0.102 |
| | DOS | -0.601 | 0.831 | 1.078 | -0.728 | -0.128 |
| | Sen2Cor | -0.479 | 0.853 | 1.081 | -0.737 | -0.012 |
| | Rayleigh | -0.695 | 0.862 | 1.115 | -0.847 | 0.066 |
| Mango tree | iCOR | -1.497 | 1.502 | 1.754 | -2.949 | -0.028 |
| | DOS | -1.759 | 1.759 | 1.975 | -4.009 | -0.042 |
| | Sen2Cor | -1.876 | 1.876 | 2.094 | -4.632 | -0.085 |
| | Rayleigh | -1.990 | 1.990 | 2.186 | -5.135 | -0.04 |
| Miombo forest | DOS | -0.065 | 0.436 | 0.542 | -0.073 | 0.075 |
| | Rayleigh | -0.351 | 0.544 | 0.694 | -0.758 | 0.053 |
| | iCOR | 0.499 | 0.636 | 0.756 | -1.086 | 0.084 |
| | Sen2Cor | 0.368 | 0.671 | 0.824 | -1.479 | 0.049 |
| Mixed crops | iCOR | -1.090 | 1.160 | 1.373 | -2.014 | -0.022 |
| | DOS | -1.299 | 1.307 | 1.534 | -2.76 | -0.046 |
| | Sen2Cor | -1.380 | 1.390 | 1.622 | -3.201 | -0.101 |
| | Rayleigh | -1.549 | 1.549 | 1.739 | -3.832 | 0.089 |
| Sweet potato | iCOR | -0.953 | 1.188 | 1.344 | -1.504 | 0.396 |
| | DOS | -1.220 | 1.349 | 1.463 | -1.967 | 0.413 |
| | Rayleigh | -1.307 | 1.488 | 1.625 | -2.661 | 0.242 |
| | Sen2Cor | -1.07 | 1.515 | 1.705 | -3.031 | 0.218 |

## IV. DISCUSSION

### A. *Performance Comparison of Atmospheric Correction Methods*

The superior performance of iCOR across most conditions can be attributed to its comprehensive atmospheric modeling framework, which explicitly accounts for aerosol optical thickness, water vapor content, and adjacency effects. This contrasts sharply with the Rayleigh correction, which addresses only molecular scattering while ignoring aerosol contributions, and DOS, which relies on simplified empirical assumptions about dark object reflectance. Sen2Cor, despite implementing the Second Simulation of Satellite Signal in the Solar Spectrum (6S) radiative transfer code, showed intermediate performance, possibly due to reliance on climatological aerosol models that may not capture local atmospheric conditions accurately.

The poor performance of Rayleigh and DOS corrections at medium and high LAI levels reflects a crucial limitation: residual atmospheric effects introduce systematic biases into surface reflectance estimates that propagate through the PROSAIL neural network inversion. The SNAP biophysical processor's neural network was trained on top-of-canopy reflectance simulations generated by coupling PROSPECT leaf optical properties with SAIL canopy radiative transfer under standard atmospheric conditions. When atmospheric correction incompletely removes atmospheric effects, the resulting reflectance spectra fall outside the training domain of the neural network, forcing extrapolation that produces unreliable LAI estimates. This explains why simpler atmospheric corrections, despite being computationally efficient, resulted in larger systematic errors and negative $R^2$ values.

### B. *Sensitivity to Canopy Density and Vegetation Structure*

The strong positive correlation between the absolute error and field LAI value ($r$ = 0.733-0.895) represents a critical finding with operational implications. This pattern indicates that retrieval accuracy deteriorates progressively as canopy density increases. At high LAI values, canopy reflectance becomes increasingly insensitive to further LAI increments due to signal saturation in the near-infrared region, a phenomenon documented extensively in radiative transfer theory. Similar sensitivity patterns were reported in [10], demonstrating that LAI estimation errors increase with both canopy cover and spectral variability.

The reversal in method ranking between low and higher LAI classes further underscores the importance of context-specific evaluation. DOS performed the best at low LAI (RMSE = 0.478) because sparse canopies expose substantial bare soil, where the DOS assumption of zero reflectance for dark pixels approximates reality more closely. However, as canopy closure increases, this assumption breaks down, and the more sophisticated atmospheric modeling of iCOR becomes advantageous.

Vegetation-type specific differences likely reflect structural and spectral characteristics that influence both atmospheric correction and radiative transfer inversion. Miombo forest's superior performance (RMSE = 0.542) may result from its relatively homogeneous canopy structure and minimal soil background influence at the 10 m pixel scale. In contrast, Mango orchards' poor performance (RMSE = 1.754) probably stems from their typically sparse planting geometry, creating significant sub-pixel heterogeneity and strong soil-vegetation contrast. These findings align with previous observations [13] documented in South African agricultural systems, where heterogeneous canopy structures similarly reduced LAI retrieval accuracy.

### C. *Spatial Variability and Environmental Factors*

The performance contrast between the ESUDER Campus zone (RMSE = 0.518-0.745) and the Machengue zone (RMSE = 1.435-1.741) points to site-specific factors influencing retrieval accuracy. Potential explanations include differences in atmospheric conditions during image acquisition, variations in soil background optical properties, terrain effects, and agricultural management practices affecting canopy uniformity. The fact that DOS achieved a positive $R^2$ of 0.194 only at ESUDER Campus zone suggests that atmospheric conditions at this location may have been particularly clear, allowing the simple DOS assumption to approximate reality more successfully. Similar site-specific variations were documented in [14] in Polish agricultural landscapes, where local environmental conditions significantly modulated retrieval performance.

The lower error variance observed at ESUDER Campus zone may also reflect more controlled agricultural management, typical of research stations, resulting in greater within-field homogeneity compared to smallholder systems at Machengue zone. This interpretation is consistent with the findings in [6], which reported lower agreement in smallholder maize fields in Mozambique, attributing reduced accuracy to field-scale heterogeneity and mixed cropping patterns that violate the assumption of spectrally homogeneous pixels.

*D. Systematic Underestimation and Negative $R^2$ Values*

The persistent negative $R^2$ values across most stratification levels and locations indicate a significant challenge: satellite-derived LAI estimates failed to explain the variance in field LAImean. This finding raises important questions about the compatibility between effective LAI measured with hemispherical photography, which integrates all photosynthetically active plant elements, including stems and branches, and the PROSAIL-based neural network retrieval, which theoretically estimates green LAI from leaf optical properties.

This scale mismatch has been recognized in [5, 15], where discrepancies between ground-based effective LAI and satellite-retrieved green LAI contributed to systematic biases. Mixed pixel effects at the 10 m spatial resolution may further degrade retrievals in heterogeneous landscapes where sub-pixel variability in soil background, vegetation structure, and canopy gaps cannot be adequately represented.

The systematic negative Bias (-0.684 to -1.185) suggests that FVC normalization alone cannot compensate for significant mismatches between field reference measurements and satellite retrieval assumptions. This underestimation may also reflect the neural network's tendency to regress toward mean training values when encountering input spectra outside its training domain, a common limitation of empirical retrieval algorithms documented in [9] regarding European rice cultivation systems.

*E. Implications for Operational Monitoring*

Despite the challenges identified, the differential performance across LAI classes offers practical guidance for operational applications. For sparse vegetation (LAI $\leq$ 1.31), DOS provides adequate accuracy (RMSE = 0.478) with minimal computational requirements, making it suitable for rapid assessments in resource-limited settings. However, for medium to dense canopies, iCOR should be prioritized despite higher computational costs, as it consistently outperformed other methods with an RMSE of 0.956 for medium LAI, and 1.929 for high LAI.

The vegetation-type specific patterns suggest that crop-specific quality flags or uncertainty estimates should accompany operational LAI products. For example, LAI retrievals over tree crops such as mango should be flagged with higher uncertainty, while estimates over relatively homogeneous crops such as maize or natural forests may be assigned higher confidence levels.

*F. Limitations and Future Research Directions*

This study's reliance on FVC normalization, while reducing sub-pixel heterogeneity effects, may have introduced unintended biases that require further investigation. The use of a single PROSAIL-based neural network algorithm precluded assessment of whether alternative LAI retrieval approaches, such as physically-based optimization inversions and machine learning with site-specific training, might better leverage improved atmospheric correction. Temporal sampling limitations, with imagery acquired at different phenological stages across sites, prevented evaluation of seasonal consistency in method performance. Future studies should implement multi-temporal validation across complete growing seasons to assess whether atmospheric correction effects vary with phenology.

The limited number of locations (two primary sites) restricts the generalization of spatial patterns. Expanded validation across diverse agro-ecological zones, soil types, and management systems would strengthen recommendations for operational deployment. Additionally, integrating higher spatial resolution imagery such as PlanetScope, Sentinel-2 super-resolution could help disentangle whether observed limitations stem primarily from atmospheric correction inadequacies or significant spatial resolution constraints.

Future research should explore hybrid retrieval frameworks that combine multiple atmospheric correction methods based on automated scene characterization (e.g., atmospheric optical thickness, terrain complexity, vegetation density). Development of local calibration procedures for PROSAIL neural networks, trained on region-specific canopy architectures and soil backgrounds, could substantially improve accuracy in tropical agricultural systems. Furthermore, implementation of crop-specific LAI algorithms that account for structural differences among vegetation types may contribute to reducing vegetation-type-dependent errors.

## V. CONCLUSIONS

The evaluation of four atmospheric correction approaches for Sentinel-2 Leaf Area Index (LAI) retrieval in heterogeneous croplands revealed that method selection significantly impacts product accuracy and should be guided by vegetation characteristics, canopy density, and local environmental conditions. Image Correction (iCOR) emerged as the most robust method overall, particularly for medium to high LAI ranges and across diverse vegetation types. Dark Object Subtraction (DOS) demonstrated comparable performance for sparse canopies but was deteriorated at higher canopy densities. Sentinel-2 Correction (Sen2Cor) and Rayleigh correction consistently underperformed, likely due to incomplete atmospheric compensation that propagated into PROSAIL neural network extrapolation errors.

The strong positive correlation between retrieval error and canopy density highlights crucial challenges for optical remote sensing in dense vegetation, where signal saturation limits LAI discrimination. Pronounced spatial variability between ESUDER Campus and Machengue zones, as well as vegetation-type-specific performance differences, highlight the complexity of operational LAI monitoring in heterogeneous agricultural landscapes. The persistent systematic underestimation and negative $R^2$ values across most conditions suggest that current PROSAIL-based neural network retrievals may require local calibration or alternative algorithms to achieve reliable LAI estimates in tropical cropland environments.

For operational applications in similar heterogeneous agricultural systems, it is proposed to: (1) prioritize iCOR atmospheric correction when computational resources permit, particularly for medium to dense canopies, (2) use DOS for rapid assessments of sparse vegetation where computational efficiency is critical, (3) stratify accuracy assessments by vegetation type and canopy density to provide appropriate uncertainty estimates;, and (4) implement quality flags to identify problematic retrievals in high LAI conditions or structurally complex vegetation types.

These findings contribute to improving satellite-based crop monitoring capabilities in data-scarce tropical regions where ground-based LAI measurements remain logistically challenging. Future efforts should focus on developing hybrid retrieval frameworks that integrate physically-based radiative transfer models with machine learning approaches, enabling better representation of canopy heterogeneity and atmospheric variability while maintaining operational efficiency for large-scale agricultural monitoring applications.

## ACKNOWLEDGMENT

The authors acknowledge ICIPE for funding through Grant No. RSIF/RA/0017, the Eduardo Mondlane University, and the field assistance provided by ESUDER staff during data collection activities. The authors also thank the European Space Agency (ESA) for providing free access to Sentinel-2 imagery through the Copernicus Open Access Hub.

## REFERENCES

[1] J. M. Chen, C. H. Menges, and S. G. Leblanc, "Global Mapping of Foliage Clumping Index Using Multi-Angular Satellite Data," *Remote Sensing of Environment*, vol. 97, no. 4, pp. 447–457, Sept. 2005, https://doi.org/10.1016/j.rse.2005.05.003.

[2] "Systematic Observation Requirements for Satellite-based Data Products for Climate, 2011 Update," Global Climate Observing System, Geneva, Switzerland, Technical Report GCOS-154, 2011.

[3] F. Baret *et al.*, "GEOV1: LAI and FAPAR Essential Climate Variables and FCOVER Global Time Series Capitalizing Over Existing Products. Part 1: Principles of Development and Production," *Remote Sensing of Environment*, vol. 137, pp. 299–309, Oct. 2013, https://doi.org/10.1016/j.rse.2012.12.027.

[4] R. B. Myneni *et al.*, "Global Products of Vegetation Leaf Area and Fraction Absorbed PAR from Year One of MODIS Data," *Remote Sensing of Environment*, vol. 83, no. 1–2, pp. 214–231, Nov. 2002, https://doi.org/10.1016/S0034-4257(02)00074-3.

[5] L. A. Brown *et al.*, "Validation of Baseline and Modified Sentinel-2 Level 2 Prototype Processor Leaf Area Index Retrievals Over the United States," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 175, pp. 71–87, May 2021, https://doi.org/10.1016/j.isprsjprs.2021.02.020.

[6] S. Mananze, I. Pôças, and M. Cunha, "Retrieval of Maize Leaf Area Index Using Hyperspectral and Multispectral Data," *Remote Sensing*, vol. 10, no. 12, Dec. 2018, Art. no. 1942, https://doi.org/10.3390/rs10121942.

[7] J. Verrelst *et al.*, "Optical Remote Sensing and the Retrieval of Terrestrial Vegetation Bio-Geophysical Properties – A Review," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 108, pp. 273–290, Oct. 2015, https://doi.org/10.1016/j.isprsjprs.2015.05.005.

[8] F. J. García-Haro *et al.*, "Derivation of Global Vegetation Biophysical Parameters from EUMETSAT Polar System," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 139, pp. 57–74, May 2018, https://doi.org/10.1016/j.isprsjprs.2018.03.005.

[9] M. Campos-Taberner *et al.*, "A Critical Comparison of Remote Sensing Leaf Area Index Estimates over Rice-Cultivated Areas: From Sentinel-2 and Landsat-7/8 to MODIS, GEOV1 and EUMETSAT Polar System," *Remote Sensing*, vol. 10, no. 5, May 2018, Art. no. 763, https://doi.org/10.3390/rs10050763.

[10] M. Weiss and F. Baret, *S2 ToolBox Level 2 Products: LAI, FAPAR, FCOVER Version 1.1*. Paris, France: European Space Agency, 2016.

[11] P. S. Nagendram, P. Satyanarayana, and P. Ravi Teja, "Mapping Paddy Cropland in Guntur District Using Machine Learning and Google Earth Engine utilizing Images from Sentinel-1 and Sentinel-2," *Engineering, Technology & Applied Science Research*, vol. 13, no. 6, pp. 12427–12432, Dec. 2023, https://doi.org/10.48084/etasr.6460.

[12] R. Khalil, M. S. Khan, Y. Hasan, N. Nacer, and S. Khan, "Supervised NDVI Composite Thresholding for Arid Region Vegetation Mapping," *Engineering, Technology & Applied Science Research*, vol. 14, no. 3, pp. 14420–14427, Jun. 2024, https://doi.org/10.48084/etasr.7202.

[13] M. Kganyago, P. Mhangara, T. Alexandridis, G. Laneve, G. Ovakoglou, and N. Mashiyi, "Validation of Sentinel-2 Leaf Area Index (LAI) Product Derived from SNAP Toolbox and its Comparison with Global LAI Products in an African Semi-Arid Agricultural Landscape," *Remote Sensing Letters*, vol. 11, no. 10, pp. 883–892, Oct. 2020, https://doi.org/10.1080/2150704X.2020.1767823.

[14] Z. Bochenek *et al.*, "Validation of the Lai Biophysical Product Derived from Sentinel-2 and Proba-V Images for Winter Wheat in Western Poland," *Geoinformation Issues*, vol. 9, no. 1, pp. 15–26, 2017.

[15] N. Djamai, R. Fernandes, M. Weiss, H. McNairn, and K. Goïta, "Validation of the Sentinel Simplified Level 2 Product Prototype Processor (SL2P) for mapping cropland biophysical variables Using Sentinel-2/MSI and Landsat-8/OLI data," *Remote Sensing of Environment*, vol. 225, pp. 416–430, May 2019, https://doi.org/10.1016/j.rse.2019.03.020.

[16] J. Louis *et al.*, "SENTINEL-2 SEN2COR: L2A Processor for Users," in *Proceedings Living Planet Symposium*, Prague, Czech Republic, May 2016, Art. no. SP-740.

[17] L. De Keukelaere *et al.*, "Atmospheric Correction of Landsat-8/OLI and Sentinel-2/MSI Data Using iCOR Algorithm: Validation for Coastal and Inland Waters," *European Journal of Remote Sensing*, vol. 51, no. 1, pp. 525–542, Jan. 2018, https://doi.org/10.1080/22797254.2018.1457937.

[18] S. Sterckx, S. Knaeps, S. Kratzer, and K. Ruddick, "SIMilarity Environment Correction (SIMEC) Applied to MERIS Data Over Inland and Coastal Waters," *Remote Sensing of Environment*, vol. 157, pp. 96–110, Feb. 2015, https://doi.org/10.1016/j.rse.2014.06.017.

[19] A. B. Ruescas and D. Müller, *Rayleigh Correction Tutorial*. Paris, France: European Space Agency, 2021.

[20] L. Congedo, "Semi-Automatic Classification Plugin: A Python Tool for the Download and Processing of Remote Sensing Images in QGIS," *Journal of Open Source Software*, vol. 6, no. 64, Aug. 2021, Art. no. 3172, https://doi.org/10.21105/joss.03172.