



Enhancing leaf area index estimation in tropical vegetation: a comparative study of multivariate linear regression and Sentinel Application Platform-derived leaf area index

Ali Yasin Ahmed^{1*}, Abebe Mohammed Ali², Nurhussen Ahmed² and Birhane Gebrehiwot³

¹Department of Geography and Environmental Studies, Jigjiga University, Jigjiga 1020, Ethiopia

²Department of Geography and Environmental Studies, Wollo University, Dessie 1145, Ethiopia

³Department of Land Administration and Surveying, Dilla University, Dilla 419, Ethiopia

ARTICLE INFO

Received July 2, 2024

Revised September 11, 2024

Accepted September 11, 2024

Published on October 29, 2024

*Corresponding author

Ali Yasin Ahmed

E-mail alexoy5050@gmail.com

Background: The leaf area index (LAI) quantifies the total one-sided green leaf area per unit of soil area, making it a crucial parameter in models that simulate carbon, nutrient, water, and energy fluxes within forest ecosystems. This study enhances LAI estimation techniques by employing a multivariate linear regression (MVLRL) approach specifically tailored to tropical vegetation. We integrated field-collected LAI data with spectral indices and multispectral bands to develop a robust predictive empirical model. The LAI estimates derived from the MVLRL approach are rigorously compared with those obtained from the Sentinel Application Platform (SNAP), a widely utilized tool for remote sensing analysis.

Results: In developing the MVLRL model, nine multispectral bands, seven vegetation indices (VIs), and two biophysical variables derived from Sentinel-2 multispectral image were tested to identify efficient predictors for LAI estimation. To determine significant multispectral bands and VIs (ensuring no multicollinearity, high coefficient of determination (R^2), low root mean square error (RMSE), and a p -value < 0.05) for the best representative model, stepwise multiple linear regression (SMLRL) was employed. Multispectral bands 7 and 8, along with the VIs soil adjusted vegetation index and normalized difference vegetation index, and the fraction of vegetation cover biophysical variable, produced superior outcomes and serve as strong predictor variables for LAI. The accuracy of the MVLRL model was validated using 17 directly measured LAI sample plots with the leave-one-out cross-validation method. The estimated LAI using the MVLRL model achieved higher accuracy, with an R^2 of 0.94, compared to the SNAP toolbox ($R^2 = 0.71$). The RMSE and bias of the MVLRL model were 0.18 and 0.006, respectively, while for SNAP-derived LAI, the RMSE and bias were 0.53 and 0.31, respectively.

Conclusions: The improved accuracy and reduced error of the MVLRL model are attributed to its adjustment for tropical vegetation types. Future research should focus on comparing the MVLRL model with other global LAI products to further validate and enhance its applicability.

Keywords: leaf area index, multispectral bands, multivariate, Sentinel-2, vegetation indices

Introduction

Leaf area index (LAI) is one of the key biophysical parameters in monitoring various ecological processes. It can be efficiently applied in monitoring vegetation photosynthesis, respiration, transpiration, soil respiration, and the energy exchange of the canopy-atmosphere (Jin and Zhang 2002; Li et al. 2023; Mwangi et al. 2018; Zhang et al. 2020). This demonstrates that LAI is applicable in many fields of study like precision agriculture, vegetation health monitor-

ing, and crop yield estimation (Ma et al. 2022; Sishodia et al. 2020; Xu et al. 2020; Yao et al. 2017). LAI also plays a key role in climate modeling, and biodiversity monitoring; this also points out that LAI is an essential climate variable (Darvishzadeh et al. 2019; Reygadas et al. 2020). It has a strong positive correlation with net primary production, aboveground biomass, and agricultural yield (Eisfelder et al. 2017; Muhe and Argaw 2022). On the other hand, in many instances, it has a negative correlation with climatic variables like surface temperature (Eisfelder 2013; Pizaña

© 2024 The Author(s) **Open Access**



This article is licensed under a Creative Commons Attribution (CC BY) 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>. The publisher of this article is The Ecological Society of Korea in collaboration with The Korean Society of Limnology.

et al. 2016).

LAI can be estimated either through direct or destructive sampling or by employing indirect methods that estimate LAI based on the degree of light penetration through the canopy (Ghebremicael et al. 2004). Although direct or destructive measurements of LAI offer precision, collecting such data is labor-intensive and time-consuming, rendering it feasible only for small-scale plots (Vafaei et al. 2021). The availability of freely accessible, high-resolution satellite datasets derived from advanced sensors such as Sentinel-2, coupled with the utilization of sophisticated open-source tools like the Sentinel Application Platform (SNAP), and the growing accessibility of analysis-ready data, presents a remarkable opportunity for the prediction of precise vegetation variables (Kganyago et al. 2020).

In remote sensing, four basic approaches can be used to retrieve or estimate biophysical variables such as LAI from satellite images. First, physical modeling uses radiative transfer models (RTM) to predict the interaction between spectral radiation and vegetation biophysical variables (Cho et al. 2008; Jacquemoud et al. 2000; Myneni 1997; Sikand and Stamnes 2012). Second, empirical or parametric modeling establishes a relationship between vegetation indices (VIs) and the ground measured biophysical variables. Third, machine learning regression algorithm or non-parametric modeling do not require an explicit relationship or specific data distribution, for-example, support vector regression (Durbha et al. 2007). Lastly, hybrid models combine elements from two or more of the aforementioned approaches (Lillesand et al. 2015; Verrelst et al. 2015).

In recent studies, RTMs are used to derive biophysical variables from satellite images (O'Hagan 2006; Verrelst et al. 2016). These RTMs utilize the full spectrum acquired by multi- to hyperspectral sensors, unlike VIs that use only some of the spectral bands (Verrelst et al. 2016). Additionally, RTMs can exploit the directional signature of multi-angle sensors (Verrelst et al. 2015). However, these models have some weaknesses, including the need for extensive parameterization, as well as requiring more intensive computational skills and time (Richter et al. 2011). On the other hand, most of the empirical models do not require advanced mathematical skills, making them easier to understand and implement (Ali et al. 2019; Magney et al. 2017).

The SNAP biophysical processor, an open-access earth observation analysis tool, estimates biophysical parameters based on two different hybrid models: a physically-based RTM and a robust machine learning algorithm (Upreti et al. 2019; Weiss and Baret 2016). The accuracy of SNAP-derived LAI from Sentinel-2 data, as well as its consistency with existing LAI products, has been studied (Bochenek et al. 2017; Campos-Taberner et al. 2018; Kganyago et al. 2020). This is crucial since precise and consistent agricultural monitoring strongly hinges on the consistency and inter-comparability of biophysical parameters (Alexan-

dridis et al. 2019).

While numerous studies (Binh et al. 2022; Brown et al. 2019; Chauhan and Lunagaria 2024; Djamai et al. 2019a; Kganyago et al. 2020; Pasqualotto et al. 2019b) have validated LAI estimates derived from Sentinel-2 imagery using the SNAP biophysical variables processor, there remains a gap in comparing these estimates with empirical models that incorporate both multispectral bands and VIs. Our study addresses this gap by developing and validating a multivariate linear regression (MVLr) model that integrates these spectral features, providing a fresh perspective on LAI estimation accuracy. This research, conducted in the Mille watershed of Ethiopia, offers a critical comparison between SNAP-derived LAI and our tailored MVLr model, aiming to enhance LAI estimation techniques, especially in tropical regions with complex vegetation types. The significance of our study lies in its ability to provide more accurate and context-specific LAI estimates, which are essential for improving applications in climate modeling, agricultural management, and forestry practices in regions with similar ecological characteristics. Hence, this study has two main objectives: (1) identify the most efficient predictors of LAI from satellite image reflectance and VIs and develop a MVLr model to estimate LAI and (2) validate and compare LAI derived from SNAP biophysical processor and MVLr model against field-measured LAI.

Methods and Materials

Description of the study area

The research was conducted in the Millie River basin, located in the Ambasel District of the South Wollo Zone in the Amhara Regional State of Ethiopia, an area predominantly characterized by tropical vegetation types (Fig. 1). Mille River basin is located between 11° 19' 44"–11° 32' 37" latitude and 39° 31' 44"–39° 40' 51" longitude. It has an elevation ranging from 1,427 to 3,635 m above sea level with an average temperature 15°C to 20°C.

The rainfall pattern is bimodal type. The main rainy season is from June to September, while the medium rainy season is from April to June. According to the Kombolcha metrological station, the mean annual rainfall of the study area ranges from 800 to 2,500 mm. The lowest minimum temperature is recorded from August to November (11°C to 12°C) and the highest maximum temperature is recorded during May and June (22°C to 30°C). The livelihood of the study area is mainly mixed agriculture i.e., crop-livestock production, which is mainly rain-fed agriculture.

According to the district statistics, more than 68% of the Mille watershed is covered by agricultural crops and vegetations, while the northern and northwestern parts are dominated by medium-height forest trees, such as *Prunus africana*, *Croton macrostachyus*, *Cordia africana*, and *Fi-*

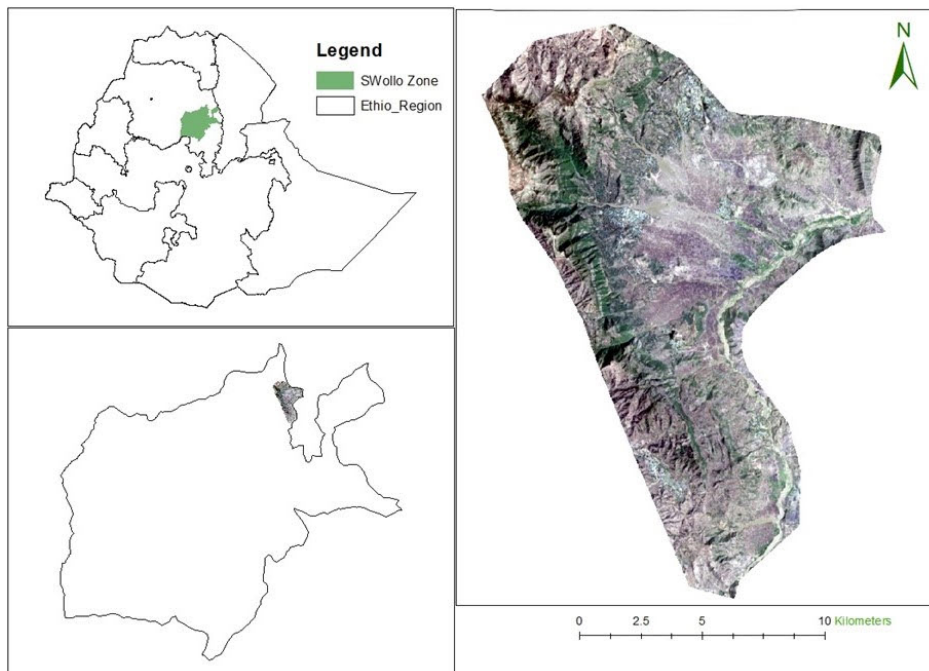


Fig. 1 Map of the Millie Watershed, located in a tropical climatic region of Ethiopia. The map features a true-color composite of the Sentinel-2 satellite image, utilizing bands 4, 3, and 2. This combination provides an accurate representation of the watershed's land cover and vegetation, effectively highlighting the key geographical and environmental features pertinent to our study.

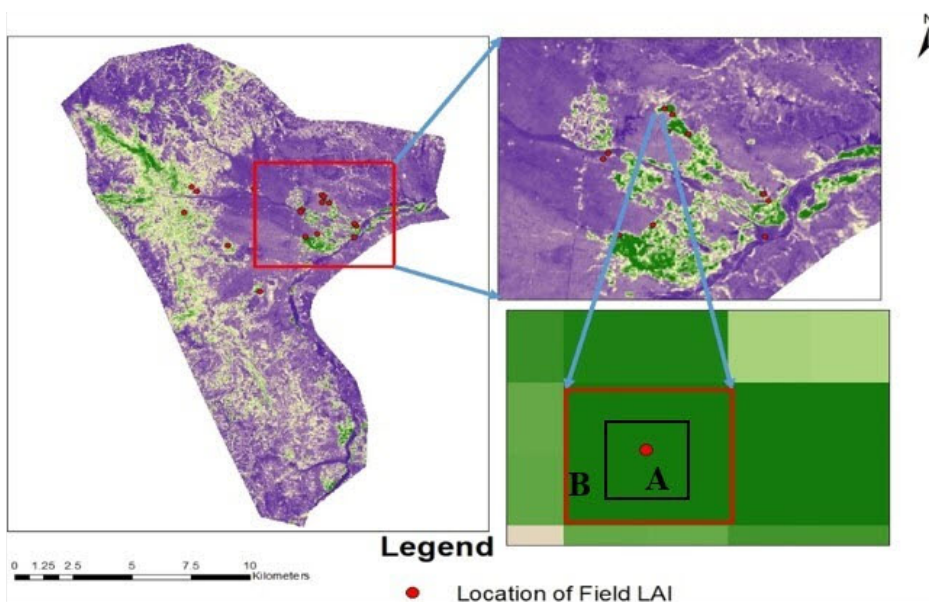


Fig. 2 Location of sample LAI distribution measured in the field. The figure shows 17 sample pixels where LAI was measured using a direct/destructive method with minimal impact on the vegetation. Letter A (black box) indicates a 10 m × 10 m area initially delineated for each sample plots to minimize disturbance during leaf area measurement using the Petiole application. Letter B (red box) marks a 20 m × 20 m area assessed based on the initial measurement in letter A to further reduce disturbance. This method provides accurate ground-truth data for validating the SNAP and MVLr models (Table 1). LAI: leaf area index; SNAP: Sentinel Application Platform; MVLr: multivariate linear regression.

cus sur. The study area is characterized by a rich diversity of vegetation and agricultural practices. Economically significant plant species, including *Coffea arabica*, *Mangifera indica*, and *Catha edulis*, play a vital role in the local economy. In addition to these perennial crops, the region also supports various seasonal agricultural crops, such as *Zea mays*, *Sorghum bicolor*, *Eragrostis tef*, and *Capsicum annum*, which are particularly abundant during the summer growing season. Vegetation coverage within the watershed varies, with some areas exhibiting dense vegetation, while others are more sparsely covered. The Mille River basin is illustrated using a Sentinel-2 true color composite in Figure 1.

Field LAI measurement

Due to the lack of indirect LAI measurement instruments in the field, the researchers employed direct measurement with minimal disturbance for 17 sample plots. Each field-measured sample plot has a 20 m × 20 m area (Fig. 2B). Before taking measurements, the researchers recorded global positioning system (GPS) coordinates for each plot. The central pixel GPS readings for all plots were verified by overlaying the coordinates with Sentinel-2 image at 20-m resolution.

During field measurement, 10 m × 10 m plot sizes were initially delineated for 17 plots (Fig. 2A). Subsequently, a minimum of five evenly distributed vegetation samples were deliberately selected based on criteria such as height,

leaf width, and representativeness. From these chosen plants, typical branches located in the middle of each plant were carefully selected for sampling, ensuring minimal disturbance and efficiency in measurement. The area of each leaf was assessed using a smartphone application called Petiole, which utilizes the smartphone's camera for individual leaf area measurement.

Before use, the camera was calibrated using a calibration pad provided by Petiole Limited Organization. Petiole, a plant leaf area meter app is a free Android app, which has been published by Petiole Limited on April 05, 2021. It is possible to measure individual leaf area, acquire total leaf area, and save data with the Petiole app. Petiole measures most of the vegetation leaf such as soybean, corn, wheat, potato, coffee, pepper, mango, and others with similar leaf structure as mentioned vegetation type. Petiole Pro, a commercial app, helps to monitor plant health, choose plant phenotype for screening, and understand plant stress faster (Confalonieri et al. 2013; Singh et al. 2021).

It can be downloaded for free from <https://cloudapks.com/download/416872/26/>. LAI field measurement was conducted on non-needle leaf plants with medium height because it is not advisable to use a mobile app in tinny leave and tall forest trees (Fang et al. 2014; Singh et al. 2021). Lastly, the LAI of each sample plot was computed using Equation 1 (Jonckheere et al. 2004; Stenberg et al. 1994). The statistics of LAI measured in the study plots ($n = 17$) is provided in Table 1.

$$LAI = \frac{L \times B \times T}{PA} \quad (1)$$

where L represents the total leaf area of the sample branches in square meters (m^2), B is the total number of branches on

the sample plants, T is the total number of trees in the sample plot, and PA denotes the total area of the sample plot in square meters (m^2).

Sentinel-2 data and preprocessing

The remotely sensed data used to estimate LAI was a Level 2A cloud-free ($< 10\%$) satellite image, downloaded from <https://scihub.copernicus.eu/>. Level-2A's main output is bottom-of-atmosphere corrected reflectance, which is geometrically and atmospherically corrected (Main-Knorn et al. 2015; Sola et al. 2018).

All 10 m spatial resolution bands were resampled to a 20 m pixel size to ensure consistency across all Sentinel-2 image bands and to match the plot size of the in-situ data. The required Sentinel-2 spectral bands for the SNAP and MVLN models were then extracted for the study area, as the SNAP toolbox requires only eight specific bands.

Methods

SNAP toolbox LAI retrieval

In this study, we used the SNAP biophysical processor to retrieve the LAI from Sentinel-2 satellite images. The process relies on a combination of the PROSAIL (a combination of PROSPECT [leaf optical properties spectra] and SAIL [scattering by arbitrarily inclined leaves]) RTM and an artificial neural network (ANN) algorithm. Specifically, the SNAP biophysical processor uses reflectance data from eight Sentinel-2 bands (B3, B4, B5, B6, B7, B8a, B11, and B12) along with angles related to the sun and satellite to simulate how light interacts with the vegetation (Xie et al. 2019). The ANN, trained on this simulated data, is then used to estimate LAI from the actual Sentinel-2 image

Table 1 Leaf area index estimated from field sample plot (pixel) measurements

Plot ID	Location (m)		Vegetation (scientific name)	Measured LAI
	Latitude	Longitude		
001	570,241	1,267,174	<i>Capsicum annum</i>	1.00
002	570,160	1,267,281	<i>Capsicum annum</i>	1.50
003	567,952	1,267,956	<i>Bare-land</i>	0.00
004	567,884	1,267,857	<i>Catha edulis</i>	0.48
005	568,876	1,268,594	<i>Mangifera indica</i>	2.40
006	568,765	1,268,686	<i>Mangifera indica</i>	1.80
007	569,097	1,268,266	<i>Mangifera indica</i>	2.00
008	566,150	1,263,929	<i>Mangifera indica</i>	1.80
009	562,926	1,267,815	<i>Coffea arabica</i>	1.85
010	568,110	1,266,612	<i>Mangifera indica</i>	2.30
011	568,590	1,266,774	<i>Mangifera indica</i>	2.00
012	570,191	1,266,590	<i>Bare-land</i>	0.00
013	563,480	1,268,868	<i>Catha edulis</i>	1.00
014	563,254	1,269,081	<i>Coffea arabica</i>	1.50
015	565,904	1,268,952	<i>Catha edulis</i>	1.18
016	568,809	1,268,310	<i>Mangifera indica</i>	2.50
017	564,805	1,266,188	<i>Coffea arabica</i>	1.40

LAI: leaf area index.

(Weiss and Baret 2016).

The ANN training data is created using an RTM, which accurately simulates canopy reflectance within Sentinel-2 bands and across various vegetation types and environmental conditions. The SNAP biophysical processor integrates both quantitative and qualitative quality indicators to evaluate the dependability of the resulting product. These indicators allow us to gauge the data’s confidence level and make informed decisions regarding its application.

This algorithm falls under the classification of “non-specific,” suggesting its ability to provide satisfactory performance across different vegetation types in terms of biophysical variables (Kamenova and Dimitrov 2020; Kganyago et al. 2020; Mourad et al. 2020; Xie et al. 2019). The SNAP toolbox includes three biophysical processors: S2_20 m, S2_10 m, and Landsat 8 (Mourad et al. 2020). In this study, we utilized the S2_20 m biophysical processor.

Vegetation indices retrieval

For this study, eighteen VIs and multispectral bands were selected to assess their effectiveness in estimating LAI for tropical vegetation. The VIs was derived from Sentinel-2 satellite imagery, following a rigorous preprocessing protocol to ensure data quality and accuracy. These indices, known for their ability to capture various aspects of vegetation health and structure, were selected for their potential to effectively assess vegetation condition (Kamenova and Dimitrov 2020; Magney et al. 2017; Xie et al. 2018).

The retrieval process involved extracting pixel values from each of the Vis’. This data was then used in subsequent regression analyses to develop a linear regression model tailored for LAI estimation in the study area. The VIs and multispectral bands tested to develop the efficient MVLR model are listed in Table 2.

Extraction of predictor variables pixel values

Pixel values for each variable derived from the Sentinel-2 image were extracted using ArcGIS 10.8 software. The geographical coordinates (latitude and longitude) of the field plots served as reference points for aligning the pixels, as illustrated in Figure 2. The extracted pixel values for each predictor variable are presented in Tables 3 and 4. These values were then exported in comma-separated values format for further correlation and regression analysis using R studio.

Multi-collinearity test

The presence of multi-collinearity between bands and VIs was tested using tolerance and variance of inflation factor (VIF) using R. Tolerance less than 0.1 and VIF greater than 10 indicates multi-collinearity between variables (Alin 2010; Daoud 2017; Joseph et al. 2012). To remove variables that have multi-collinearity with other variables, we have used the *olsrr* package in R studio (Kim 2019). This test is important to exclude variables which has the same impact on the dependent variable.

Table 2 Sentinel-2 multispectral bands and vegetation indices tested for the estimation of leaf area index

Band No.	Function	Central wavelength (nm)	Spatial resolution (m)
B3	Green	560	10
B4	Red	665	10
B5	Red-edge	705	20
B6	Red-edge	749	20
B7	Red-edge	783	20
B8	Near infrared	842	10
B8a	Near infrared	865	20
B11	Short wave IR	1,610	20
B12	Short wave IR	2,190	20

VIs	Vegetation indices	Formula
IRECI	Inverted red-edge chlorophyll index	$(\text{Band } 7 - \text{Band } 4) / (\text{Band } 5 / \text{Band } 6)$
NDVI	Normalized difference vegetation index	$(\text{Band } 8 - \text{Band } 4) / (\text{Band } 8 + \text{Band } 4)$
TNDVI	Transformed normalized difference vegetation index	$\frac{1}{2} \left(\frac{\text{NIR} - \text{RED}}{\text{NIR} + \text{RED} + 1} + 1 \right)$
NDVI45	Normalized difference vegetation index with bands 4 and 5	$(\text{Band } 5 - \text{Band } 4) / (\text{Band } 5 + \text{Band } 4)$
SAVI	Soil adjusted vegetation index	$[(\text{NIR} - \text{Red}) / (\text{NIR} + \text{Red} + 0.5)] \times (1 + 0.5)$
REP	Red-edge position	$705 + 35 \times [(B4 + B7) / 2 - B5] / (B6 - B5)$
MSAVI	Modified soil-adjusted vegetation index	$\frac{22 \times \text{NIR} + 1 - \sqrt{(2 \times \text{NIR} + 1)^2 - 8 \times (\text{NIR} - \text{RED})}}{2}$

BPVs	Biophysical variables
FAPAR	Fraction of absorbed photosynthetically active radiation
FVC	Fraction of vegetation cover

IR: infrared.

Table 3 Pixel values extracted from selected bands of Sentinel-2 image

Plot ID	Location (m)		B3	B4	B5	B6	B7	B8	B8a	B11	B12
	Latitude	Longitude									
001	570,241	1,267,174	0.188	0.181	0.241	0.349	0.381	0.388	0.393	0.300	0.237
002	570,160	1,267,281	0.185	0.180	0.238	0.365	0.401	0.389	0.419	0.288	0.223
003	567,952	1,267,956	0.212	0.228	0.239	0.252	0.257	0.250	0.265	0.250	0.226
004	567,884	1,267,857	0.176	0.192	0.213	0.247	0.260	0.278	0.274	0.270	0.244
005	568,876	1,268,594	0.175	0.157	0.223	0.396	0.439	0.446	0.455	0.281	0.209
006	568,765	1,268,686	0.192	0.169	0.233	0.361	0.393	0.438	0.414	0.283	0.219
007	569,097	1,268,266	0.174	0.155	0.213	0.404	0.470	0.490	0.491	0.274	0.197
008	566,150	1,263,929	0.169	0.153	0.204	0.385	0.440	0.462	0.465	0.281	0.199
009	562,926	1,267,815	0.162	0.166	0.209	0.298	0.338	0.368	0.359	0.258	0.198
010	568,110	1,266,612	0.180	0.159	0.231	0.430	0.479	0.505	0.502	0.296	0.215
011	568,590	1,266,774	0.191	0.174	0.245	0.377	0.411	0.442	0.444	0.315	0.234
012	570,191	1,266,590	0.217	0.234	0.247	0.253	0.257	0.246	0.257	0.253	0.233
013	563,480	1,268,868	0.170	0.169	0.226	0.311	0.337	0.332	0.358	0.298	0.222
014	563,254	1,269,081	0.171	0.173	0.215	0.304	0.331	0.372	0.353	0.282	0.214
015	565,904	1,268,952	0.168	0.169	0.229	0.330	0.358	0.361	0.378	0.292	0.227
016	568,809	1,268,310	0.199	0.168	0.216	0.355	0.398	0.480	0.408	0.266	0.211
017	564,805	1,266,188	0.162	0.147	0.206	0.380	0.436	0.431	0.465	0.273	0.196

The abbreviation is shown in Table 2.

Table 4 Pixel values of vegetation indices extracted from Sentinel-2 image

Plot ID	Location (m)		SAVI	REP	NDVI	NDVI45	MSAVI	IRECI	FVC	FAPAR
	Latitude	Longitude								
001	570,241	1,267,174	0.359	0.719	0.554	0.243	0.316	0.343	0.438	0.432
002	570,160	1,267,281	0.353	0.719	0.555	0.243	0.309	0.426	0.535	0.531
003	567,952	1,267,956	0.048	0.714	0.091	0.046	0.039	0.042	0.119	0.130
004	567,884	1,267,857	0.171	0.719	0.321	0.102	0.140	0.092	0.164	0.187
005	568,876	1,268,594	0.447	0.720	0.687	0.335	0.401	0.647	0.605	0.602
006	568,765	1,268,686	0.401	0.718	0.617	0.288	0.356	0.433	0.521	0.508
007	569,097	1,268,266	0.517	0.724	0.729	0.293	0.479	0.906	0.711	0.727
008	566,150	1,263,929	0.487	0.723	0.722	0.298	0.444	0.786	0.636	0.647
009	562,926	1,267,815	0.347	0.723	0.589	0.224	0.298	0.360	0.463	0.497
010	568,110	1,266,612	0.567	0.723	0.761	0.370	0.537	1.016	0.750	0.753
011	568,590	1,266,774	0.560	0.723	0.751	0.343	0.528	1.037	0.742	0.743
012	570,191	1,266,590	0.044	0.699	0.081	0.051	0.036	0.027	0.088	0.082
013	563,480	1,268,868	0.296	0.718	0.523	0.257	0.250	0.275	0.397	0.400
014	563,254	1,269,081	0.332	0.721	0.543	0.210	0.287	0.281	0.391	0.413
015	565,904	1,268,952	0.337	0.719	0.566	0.260	0.290	0.331	0.432	0.452
016	568,809	1,268,310	0.479	0.724	0.688	0.253	0.438	0.706	0.636	0.647
017	564,805	1,266,188	0.460	0.723	0.710	0.337	0.414	0.676	0.615	0.636

The abbreviation is shown in Table 2.

Multivariate linear regression model

In this study, we employed a MVLR model to estimate the LAI using Sentinel-2 multispectral bands and VIs. The MVLR model was developed and analysed using R Studio, chosen for its computational efficiency and its ability to handle multiple predictor variables simultaneously, which is crucial for accurately estimating vegetation characteristics from remotely sensed data.

We developed empirical models by establishing relationships between LAI and the selected VIs and multispectral bands. The Sentinel-2 bands and VIs used in this analysis are listed in Table 2. These variables were chosen based on their relevance in previous studies, which have demon-

strated that incorporating multiple VIs can significantly enhance prediction accuracy (Ali et al. 2019; Middinti et al. 2017; Zhang et al. 2021).

The process involved first extracting the spectral bands and VIs from the Sentinel-2 imagery. These extracted features were then input into the MVLR model within R Studio to establish a predictive relationship with the ground-truth LAI measurements. The regression coefficients were derived to optimize the model for LAI estimation, ensuring that the model provided accurate and reliable outputs for the specific tropical vegetation characteristics of the Mille watershed.

Validation of the model

The estimated LAI was passed through a validation process using direct/destructive field-measured LAI values. Accuracy assessment of all individual VI models and MVLR derived model-derived LAI was executed using a cross-validation technique. Due to the limited sample size of field-measured LAI values, it was not feasible to divide the data into separate training and testing datasets for the empirical models. Therefore, the training and the testing data were the whole data sets i.e., seventeen sample plots.

The best cross-validation method to use the whole data sets for training and testing is the leave-one-out cross-validation (LOOCV) (Cawley 2006). To validate the empirical model using the LOOCV method, we used the Caret package in R software. The root mean square error (RMSE), bias and coefficient of determination (R^2) were used to validate the MVLR model using the LOOCV technique. The corresponding estimated LAI from SNAP was validated with the in-situ measured LAI. The highest R^2 , lowest RMSE, and a bias close to zero were considered to be the best-estimated value of the derived LAI.

$$R^2 = 1 - \frac{\sum(y_i - y'_i)^2}{\sum(y_i - \bar{y})^2} \tag{2}$$

$$RMSE = \sqrt{\frac{\sum(y_i - y'_i)^2}{n}} \times 100 \tag{3}$$

$$Bias = \frac{\sum(y_i - y'_i)}{n} \tag{4}$$

where y_i and y'_i are the actual and predicted values for sample i and n is the number of samples considered.

Results

Relationship between multispectral bands and measured LAI

The response of LAI to Sentinel-2 bands reflectance was assessed by fitting a linear regression model. The relationship between measured LAI and Sentinel-2 band reflectance was examined before applying VIs. As shown in Figure 3, the reflectance of several Sentinel-2 bands exhibits a statistically significant correlation with LAI variation.

Relationship between VIs and measured LAI

The result of the correlation analysis between LAI measured at the field and the VIs revealed that there is a strong association between the observed LAI and most of the VIs with R^2 values ranging from 0.54 to 0.83. Among the predictor VIs, transformed normalized difference vegetation index ($R^2 = 0.82$), soil adjusted vegetation index (SAVI) ($R^2 = 0.83$), red-edge position ($R^2 = 0.54$), normalized difference vegetation index (NDVI) ($R^2 = 0.83$), NDVI with bands 4 and 5 ($R^2 = 0.72$), and inverted red-edge chlorophyll index (IRECI) ($R^2 = 0.68$) were strongly correlated with LAI (Fig. 4). Since almost all VIs were similarly correlated with LAI, we have investigated the multi-collinearity of predictors for both VIs and multispectral bands of Sentinel-2 image to exclude predictor variables that have multi-collinearity with the other variables. In addition to performing a multicollinearity test, we used stepwise multiple linear regression (SMLR) to identify significant single bands and VIs that exhibit high R^2 , low RMSE, and a p -value < 0.05 for the best representative model.

The performance of VIs in estimating LAI from Sentinel-2 bottom-of-canopy reflectance was validated with seventeen field LAI measurements. Based on the fitted linear

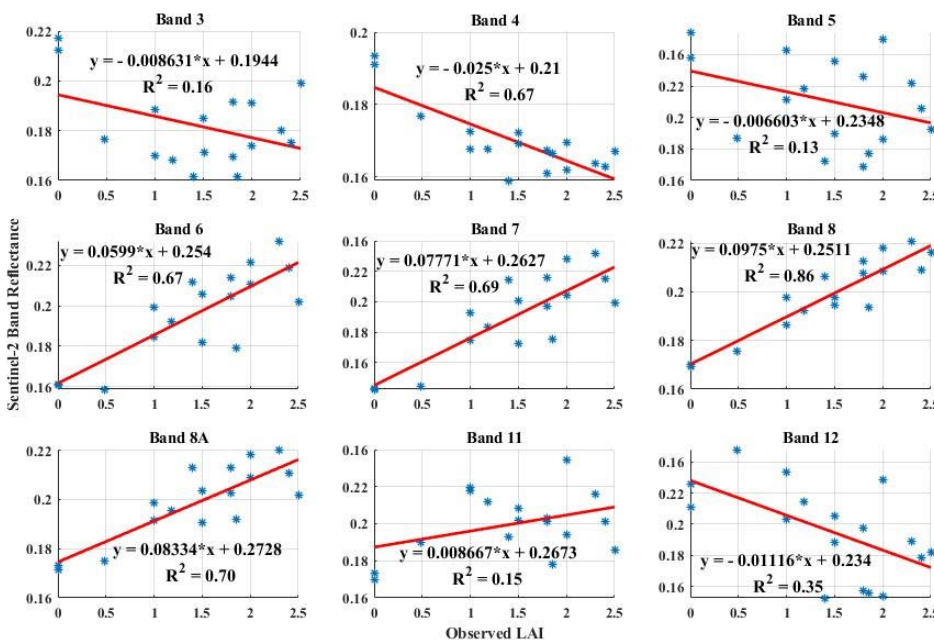


Fig. 3 Association between observed LAI and Sentinel-2 band reflectance. Among the red-edge bands, band 6 ($R^2 = 0.67$) and 7 ($R^2 = 0.69$), the near-infrared band 8 ($R^2 = 0.86$) and 8a ($R^2 = 0.70$), and band 4 ($R^2 = 0.67$) shown a strong correlation. From short wave IR spectrum band 12 has a medium association with observed LAI with an R^2 value of 0.35. Bands 3, 5, and 11 show a weak correlation with LAI measured at the field. Bands 3, 4, 5, and 12 show a negative association whereas bands from 6–11 show a positive relationship with that of observed LAI. LAI: leaf area index; IR: infrared.

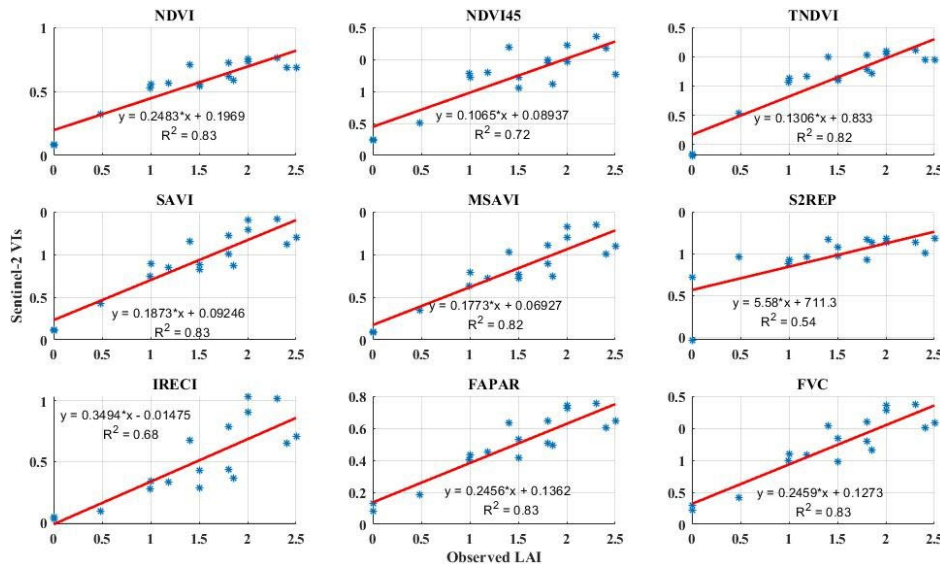


Fig. 4 Association between observed LAI and VIs computed from Sentinel-2 imagery. LAI: leaf area index; VIs: vegetation indices. The abbreviation is shown in Table 2.

Table 5 The performance of vegetation indices and biophysical variables to estimate leaf area index

No	VIs	R ²	RMSE	Bias
1	IRECI	0.59	0.48	-0.00002
2	MSAVI	0.70	0.41	-0.00009
3	NDVI	0.69	0.41	-0.000029
4	NDVI45	0.68	0.43	-0.000049
5	S2REP	0.36	0.81	-0.006
6	SAVI	0.72	0.39	-0.000028
7	TNDVI	0.63	0.45	0.00006
8	FVC	0.73	0.39	-0.000029
9	FAPAR	0.73	0.39	-0.000029
10	MVLR	0.94	0.19	0.0006
11	SNAP	0.71	0.53	0.31

The abbreviation is shown in Table 2. MVLR: multivariate linear regression; SNAP: Sentinel Application Platform.

regression (Fig. 4), the LAI of the study area was estimated. Accuracy and correlation of the VIs models in contrast to in-situ LAI were validated using LOOCV (R²_{cv}), and RMSE_{cv}. VIs models with higher accuracy include SAVI, MSAVI, NDVI, and TNDVI. Relative to the other VIs, REP has a lower correlation and accuracy (Table 5).

Selection of efficient predictors for MVLR model

When fitting a linear regression model to multispectral bands and VIs for estimating the LAI of the study area, we found that many bands and VIs are strongly correlated with the observed LAI. However, most of the tested predictors have a multi-collinearity. Then, only two bands and three VIs were selected and the remaining variables were excluded from the MVLR model because of lower R² and the presence of multi-collinearity. Table 6 shows correlation coefficient matrix between Sentinel-2 multispectral bands and VIs derived from Sentinel-2 imagery. In addition to performing a multicollinearity test, we used SMLR to identify significant single bands and VIs that exhibit high R², low RMSE, and a p-value < 0.05 for the best repre-

sentative model.

Table 7 presents the variables selected for developing the MVLR model, including specific bands and VIs, along with their statistical parameters. The variables listed are band 7, band 8, SAVI, NDVI, and fraction of vegetation cover (FVC). The intercept of the model is -0.9812. The overall model summary includes the R², adjusted R², and the model's p-value, indicating a strong fit with R² of 0.94, adjusted R² of 0.91, and a significant p-value of 3.313e-06.

Therefore, using the selected variables' correlation with field-measured LAI, and the intercept and slope of the regression, the MVLR model was developed (Equation 5).

$$\begin{aligned}
 \text{LAI} = & (-10.52 \times B7) + (13.1632 \times B8) - \\
 & (8.8494 \times \text{SAVI}) + (3.6370 \times \text{NDVI}) + \\
 & (4.9822 \times \text{FVC}) - 0.9812 \quad (5)
 \end{aligned}$$

Validation and comparison of MVLR model and SNAP toolbox derived LAI

The correlation result demonstrated that the selected VIs showed an excellent performance in predicting the LAI of

Table 6 Correlation between Sentinel-2 multispectral bands and vegetation indices derived from Sentinel-2 imagery

Variables	B7	B8	B8a	SAVI	NDVI	NDVI45	MSAVI	IRECI	FVC	FAPAR
B7	1.00									
B8	0.90	1.00								
B8a	0.98	0.90	1.00							
SAVI	0.86	0.90	0.90	1.00						
NDVI	0.79	0.87	0.85	0.88	1.00					
NDVI45	0.85	0.79	0.88	0.90	0.92	1.00				
MSAVI	0.86	0.92	0.90	0.98	0.92	0.88	1.00			
IRECI	0.85	0.83	0.83	0.87	0.74	0.72	0.90	1.00		
FVC	0.90	0.92	0.92	0.86	0.90	0.88	0.96	0.90	1.00	
FAPAR	0.88	0.92	0.92	0.94	0.92	0.86	0.96	0.90	0.98	1.00

It shows correlation coefficient matrix. The abbreviation is shown in Table 2.

Table 7 Statistical parameters of selected bands and vegetation indices for multivariate linear regression model development

Predictors	R ²	Estimate	Standard error	t-value	p-value (> t)
(Intercept)	-	-0.9812	0.7583	-1.294	0.22220
B8	0.72	13.1632	3.5448	3.713	0.00342
B7	0.68	-10.5214	3.3703	-3.122	0.00972
SAVI	0.72	-8.8494	4.2009	-2.107	0.05893
FVC	0.73	4.9822	2.3106	2.156	0.05405
NDVI	0.69	3.6370	1.6095	2.260	0.04511
Model summary					
Model type	R ²	Adjusted R ²	F-statistics	p-value	
MVLR	0.94	0.91	32.13	3.313e-06	

The abbreviation is shown in Table 2. MVLR: multivariate linear regression.

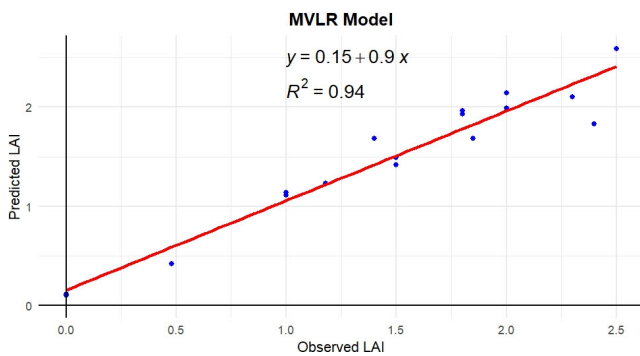


Fig. 5 Validation of MVLR model-derived LAI using ground-measured LAI. MVLR: multivariate linear regression; LAI: leaf area index.

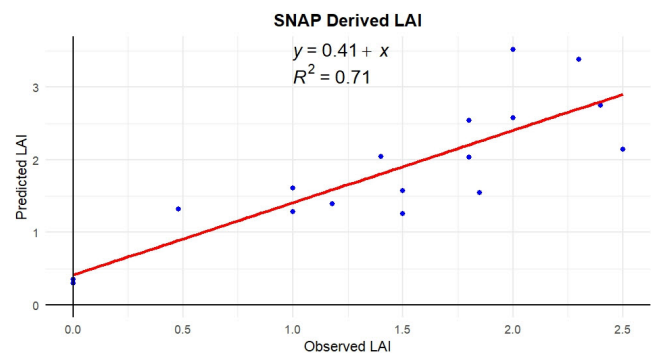


Fig. 6 Validation of SNAP toolbox-derived LAI using ground-measured LAI. SNAP: Sentinel Application Platform; LAI: leaf area index.

the study area. Even though each predictor performs well individually, the combination of band 7, band 12, IRECI, and SAVI outperforms all of them. The MVLR model was cross-validated with 17 field-measured LAI using the LOOCV technique. The RMSE and the R² show a superior performance than the SNAP-derived LAI. The developed MVLR model has an R² of 0.94 (Fig. 5). In contrast, the SNAP toolbox-derived LAI achieves an R² of 0.71 (Fig. 6). Tables 8 and 9 provide a detailed comparison of observed and predicted LAI values from both the MVLR model and

SNAP toolbox across various plot locations.

With a 95% confidence interval, LAI derived from the SNAP toolbox and the MVLR model exhibit similar outliers. RMSE for the SNAP and MVLR model-derived LAI were 0.53 and 0.19, respectively. When investigating the bias of both models, the SNAP toolbox had a bias of 0.31, whereas the MVLR model had a bias of 0.0006. The positive bias indicates overestimation in the model (Fig. 7).

Discussion

In this study, we investigated the efficacy of two distinct methodologies—MVLRL model and SNAP-derived LAI—for estimating LAI in tropical vegetation. Our results demonstrated that both methods could effectively estimate LAI; however, there were notable differences in their accuracy and reliability under varying conditions. The availability of fine Sentinel-2 bands in the red-edge region of the electromagnetic spectrum, combined with their high spatial resolution, facilitates better-quality and simpler calculations of LAI using multispectral bands and VIs. This enhanced capability allows for more accurate and detailed assessments of vegetation characteristics, which are crucial for various ecological and environmental applications.

While some VIs saturate when the LAI increases (Baret and Guyot 1991; Pasqualotto et al. 2019a), this study demonstrates that such saturation does not impact our estimates. The maximum LAI of 3.52 observed in the study area, though likely overestimated by the SNAP toolbox, falls below the saturation threshold (Baret and Guyot 1991). This suggests that the selected VIs and multispectral bands effectively capture LAI variations within the observed range, ensuring reliable and accurate measurements.

Table 8 Comparison of multivariate linear regression model and Sentinel Application Platform derived leaf area index

Models	R ²	Bias	RMSE
MVLRM derived LAI	0.94	0.0006	0.19
SNAP toolbox derived LAI	0.71	0.31	0.53

RMSE: root mean square error; MVLRM: multivariate linear regression model; LAI: leaf area index; SNAP: Sentinel Application Platform.

The results of the MVLRL model demonstrate the effectiveness of empirical models that utilize combinations of multispectral bands and VIs for accurately estimating green LAI. Such models offer simplicity and speed. The findings show that LAI strongly correlates with most of the multispectral bands and VIs derived from Sentinel-2 satellite image. Each of the tested bands and VIs has a statistically significant correlation with the field-measured LAI. However, the presence of multicollinearity among several of the tested variables complicates the development of the MVLRL model.

Predictor variables with multi-collinearity and weak association were excluded as they cause high variance in regression analysis (Muhe and Argaw 2022). The green band, band 5 from the red-edge region, band 11 and 12 from the shortwave infrared region exhibit a weak association with

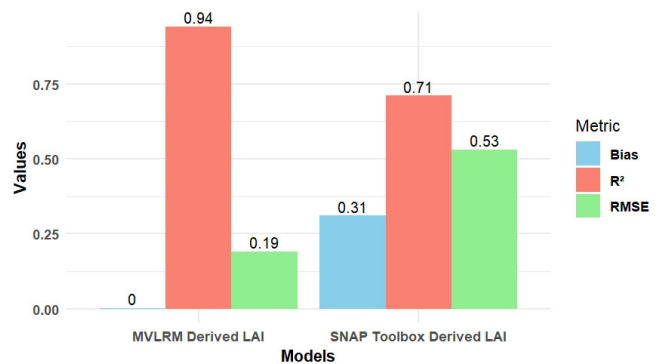


Fig. 7 Comparison of MVLRL model and SNAP toolbox: The figure illustrates the comparative analysis of LAI values derived from the MVLRL model and SNAP Toolbox methods. The graph highlights the variations in metrics and the relative performance of each method. MVLRL: multivariate linear regression; SNAP: Sentinel Application Platform; LAI: leaf area index.

Table 9 Comparison of observed LAI with predicted LAI values from the multivariate linear regression model and SNAP toolbox

Plot ID	Location (m)		Observed LAI	Predicted LAI (MVLRL) model	Predicted LAI (SNAP toolbox)
	Latitude	Longitude			
001	570,241	1,267,174	1.00	1.14	1.28
002	570,160	1,267,281	1.50	1.49	1.57
003	567,952	1,267,956	0.00	0.11	0.30
004	567,884	1,267,857	0.48	0.42	0.93
005	568,876	1,268,594	2.40	1.83	2.45
006	568,765	1,268,686	1.80	1.93	2.54
007	569,097	1,268,266	2.00	2.14	3.52
008	566,150	1,263,929	1.80	1.96	2.03
009	562,926	1,267,815	1.85	1.68	1.55
010	568,110	1,266,612	2.30	2.10	2.68
011	568,590	1,266,774	2.00	1.99	2.58
012	570,191	1,266,590	0.00	0.10	0.20
013	563,480	1,268,868	1.00	1.11	1.51
014	563,254	1,269,081	1.50	1.42	1.26
015	565,904	1,268,952	1.18	1.23	1.39
016	568,809	1,268,310	2.50	2.59	2.14
017	564,805	1,266,188	1.40	1.68	2.04

LAI: leaf area index; MVLRL: multivariate linear regression; SNAP: Sentinel Application Platform.

field-measured LAI. The combination of the above-listed bands is important in leaf dry matter content investigation rather than green LAI (Ali et al. 2019).

The estimation of LAI demonstrates promising accuracy through the utilization of individual reflectance and VIs. While each VI independently exhibits a satisfactory level of predictive capability for LAI, the incorporation of multiple VIs significantly enhances the prediction accuracy. This highlights the importance of a multifaceted approach that leverages diverse VIs to enhance LAI estimation accuracy (Middinti et al. 2017).

Multi-collinearity between the variables, assessed through tolerance and variance inflation factors, is a critical consideration in this analysis. The strong associations observed among most Sentinel-2 bands and VIs suggest that the formulation of commonly used VIs, which often rely on similar spectral bands, inherently contributes to this multi-collinearity. Specifically, the high correlations found between bands 4 and 6, as well as bands 8 and 8a, and their associations with certain VIs, are likely due to the similar reflectance characteristics of vegetation in the red and red-edge (bands 4, 6, and 7) and NIR (bands 8 and 8a) regions. These findings underscore the importance of carefully selecting input variables to minimize redundancy and enhance model performance (Zheng and Moskal 2009).

Based on the association of spectral bands with actual measured LAI and their multicollinearity statistics, the importance rankings of spectral bands for predicting LAI were as follows: NIR (band 8), red-edge (bands 6 and 7), red (band 4), SWIR (band 12), and NIR (band 8a). These results align with the findings of Chrysafis et al. (2020), who investigated LAI retrieval in a mixed Mediterranean forest area using multispectral bands derived from Sentinel-2 imagery.

The combination of band 7, band 8, SAVI, NDVI, and FVC achieved excellent accuracy in predicting LAI (Fig. 5). This result agrees with the findings of Middinti et al. (2017), who demonstrated that utilizing multiple VIs yields a more efficient and accurate prediction of LAI compared to a single VI. The enhanced accuracy can be attributed to the complementary information provided by different spectral bands and indices, which capture various aspects of vegetation characteristics and reduce the impact of individual index limitations. By integrating multiple VIs, the model benefits from a broader and more nuanced dataset, leading to improved predictive performance and robustness in LAI estimation. This multi-faceted approach allows for better handling of the complexities and variations within the vegetation canopy, thus providing a more reliable assessment of LAI.

The analysis confirms that combining multiple spectral bands and VIs significantly enhances LAI prediction accuracy. The superior performance of the MVLR model, as evidenced by the high R^2_{cv} of 0.94, emphasizes the effec-

tiveness of integrating diverse independent variables. This model's high accuracy can be attributed to the complementary information captured by different bands and indices, which provide a more comprehensive representation of the vegetation canopy. This result is highly consistent with the findings of Middinti et al. (2017), who predicted LAI using reflectance and VIs from Landsat 8 OLI (Operational Land Imager) in Indian tropical forests ($R^2 = 0.83$, RMSE = 0.78), though our study achieved a slightly higher R^2 and lower RMSE.

In our study, we validate the SNAP-derived LAI with an R^2 of 0.71, a bias of 0.31, and an RMSE of 0.53. These results demonstrate a good correlation between observed and predicted LAI values, aligning well with findings from previous studies. For instance, our R^2 value closely matches that of Djamaï et al. (2019b), who reported an R^2 of 0.70, though our RMSE (0.53) is significantly lower than their RMSE of 0.98, indicating better performance in our case. Compared to Chauhan and Lunagaria (2024), who also achieved an R^2 of 0.70 but with higher bias, our study shows improved accuracy with a lower bias (0.31). Moreover, while Binh et al. (2022) reported a lower R^2 of 0.45 and a higher RMSE of 2.19, highlighting challenges in LAI estimation for mangroves, our study's higher R^2 and lower RMSE demonstrate better reliability and precision in LAI retrieval. Similarly, our findings surpass those of Brown et al. (2019), who reported an R^2 of 0.54 and an RMSE of 1.55 in forest environments. Lastly, our results are consistent with Pasqualotto et al. (2019b), who found an R^2 greater than 0.70 and an RMSE below 0.86, further supporting the strength of LAI retrieval using SNAP in tropical regions.

However, the comparative analysis of R^2 , RMSE, and bias underscores the robustness of the MVLR model in predicting LAI. The significantly lower RMSE (0.19) of the MVLR model, compared to the SNAP toolbox's RMSE (0.53), underscores the superior precision of the MVLR model's predictions. Moreover, the minimal bias observed in the MVLR model (0.0006) suggests a reduction in systematic errors, thereby enhancing the reliability of its LAI estimates. In contrast, the positive bias observed in the SNAP-derived LAI (0.31) indicates a tendency towards overestimation, suggesting that while the SNAP toolbox is useful, it may benefit from further calibration or adjustment for specific applications. This finding supports the conclusions of Kganyago et al. (2020), who validated SNAP-derived LAI using field-measured LAI data for two different crop types—sunflower and maize. Their study identified overestimation at different growth stages of these crops, suggesting that SNAP's performance is contingent upon both crop type and growth stage, which is consistent with our findings. This indicates that the VIs and multispectral bands-based MVLR model can predict LAI with greater accuracy than the SNAP biophysical processor.

According to Kganyago et al. (2020), the overestimation

of LAI and major biases perceived from SNAP-derived LAI may be as a result of prior rules made from global LAI retrieval models which cannot be adjusted to local environments. This overestimation can impact the interpretation of vegetation health and density, making it crucial to consider the model biases when using these tools for ecological and environmental research. The findings emphasize the importance of model validation and the integration of multiple data sources to achieve accurate and reliable LAI predictions. However, the performance of SNAP-derived LAI, in this study, is slightly better than the Kganyago et al. (2020) research findings ($R^2 = 0.6$ to 0.70).

The superior accuracy of the MVLR model in our study is attributed to its adjustment for the specific vegetation types found in the study area. This tailored approach enables more precise estimation of LAI, particularly in the context of tropical climatic regions where diverse vegetation types and seasonal variability are prevalent. In contrast, the lower accuracy of the SNAP-derived LAI highlights the challenges of applying global models to local environments without customization. The effectiveness of the estimated LAI extends beyond mere accuracy; it plays a crucial role in assessing vegetation health, monitoring ecosystem dynamics, and informing sustainable agricultural practices in tropical regions. To further enhance the accuracy of the MVLR model, incorporating additional field-measured LAI data would be beneficial, providing a stronger foundation for training the model and improving its predictive capabilities across various tropical landscapes.

Conclusions

This study provides a comprehensive comparison between the MVLR model and the SNAP biophysical processor for estimating LAI in tropical vegetation. The findings demonstrate that while both methods are capable of estimating LAI, the MVLR model offers superior accuracy and reliability, particularly due to its incorporation of multispectral bands and VIs tailored to the specific characteristics of tropical vegetation. The MVLR model's high R^2 , minimal bias, and lower RMSE compared to SNAP-derived LAI underscore its effectiveness in providing precise LAI estimates, which are critical for applications in ecological monitoring and sustainable agricultural practices. The study also highlights the limitations of using global models like SNAP in local environments, where adjustments for specific vegetation types and climatic conditions are necessary. The observed overestimation in SNAP-derived LAI, aligned with findings from other studies, suggests that further calibration or adjustment is needed to improve its accuracy for local applications.

Overall, the MVLR model has demonstrated superior

performance in estimating LAI in tropical region, highlighting its potential as a robust and reliable tool. Future research should focus on incorporating additional field-measured LAI data to further refine the model and expand its applicability across diverse tropical landscapes. Additionally, comparing the MVLR model with other global LAI products will be crucial for further validation and to enhance its effectiveness in diverse environments.

Abbreviations

LAI: Leaf area index

MVLR: Multivariate linear regression

SNAP: Sentinel Application Platform

VIs: Vegetation indices

RMSE: Root mean square error

SMLR: Stepwise multiple linear regression

SAVI: Soil adjusted vegetation index

NDVI: Normalized difference vegetation index

RTM: Radiative transfer models

ANN: Artificial neural network

VIF: Variance of inflation factor

LOOCV: Leave-one-out cross-validation

IRECI: Inverted red-edge chlorophyll index

FVC: Fraction of vegetation cover

Acknowledgements

We are thankful to the members of the Department of Geography and Environmental Studies at Wollo University who assisted with field data collection. We also extend our gratitude to the anonymous reviewers and the academic editor for their valuable and insightful feedback.

Authors' contributions

AYA, AMA, and NA designed methodology, AYA collected field data. AYA, and AMA conceived the study, and designed the research framework, AYA performed statistical analysis and prepared the manuscript, AMA, and NA provided supervision throughout the study, AYA writing original draft, NA, AMA, BG Validation, AMA, NA, and BG reviewing and editing. All authors read and approved the final manuscript.

Funding

Not applicable.

Availability of data and materials

The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

References

- Alexandridis TK, Ovakoglou G, Clevers JGPW. Relationship between MODIS EVI and LAI across time and space. *Geocarto Int.* 2019; 35(13):1385-99. <https://doi.org/10.1080/10106049.2019.1573928>.
- Ali AM, Darvishzadeh R, Shahi KR, Skidmore A. Validating the Predictive power of statistical models in retrieving leaf dry matter content of a coastal wetland from a sentinel-2 image. *Remote Sens.* 2019;11(16):1936. <https://doi.org/10.3390/rs11161936>.
- Alin A. Multicollinearity. *Wiley Interdisciplinary Reviews: Computational Statistics.* 2010;2(3):370-4. <https://doi.org/10.1002/wics.84>.
- Baret F, Guyot G. Potentials and limits of vegetation indices for LAI and APAR assessment. *Remote Sens Environ.* 1991;35(2-3):161-73. [https://doi.org/10.1016/0034-4257\(91\)90009-U](https://doi.org/10.1016/0034-4257(91)90009-U).
- Binh NA, Hauser LT, Hoa PV, Thao GTP, An NN, Nhut HS, et al. Quantifying mangrove leaf area index from Sentinel-2 imagery using hybrid models and active learning. *Int J Remote Sens.* 2022;43(15-16):5636-57. <https://doi.org/10.1080/01431161.2021.2024912>.
- Bochenek Z, Dąbrowska-Zielińska K, Gurdak R, Niro F, Bartold M, Grzybowski P. Validation of the LAI biophysical product derived from Sentinel-2 and Proba-V images for winter wheat in western Poland. *Geoinf Issues.* 2017;9(1):15-26.
- Brown LA, Ogutu BO, Dash J. Estimating forest leaf area index and canopy chlorophyll content with sentinel-2: an evaluation of two hybrid retrieval algorithms. *Remote Sens.* 2019;11(15):1752. <https://doi.org/10.3390/rs11151752>.
- Campos-Taberner M, García-Haro FJ, Busetto L, Ranghetti L, Martínez B, Gilabert MA, 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 Sens.* 2018;10(5):763. <https://doi.org/10.3390/rs10050763>.
- Cawley GC. Leave-one-out cross-validation based model selection criteria for weighted LS-SVMs. Paper presented at: The 2006 IEEE International Joint Conference on Neural Network Proceedings; 2006 Jul 16-21; Vancouver, BC, Canada. Vancouver: IEEE, 2006.
- Chauhan KK, Lunagaria MM. Validation of biophysical parameters of wheat retrieved from sentinel-2 data using in-situ measurements at Anand Region of India. *Int J Environ Climate Ch.* 2024;14(1):626-38. <https://doi.org/10.9734/ijec/2024/v14i13878>.
- Cho MA, Skidmore AK, Atzberger C. Towards red-edge positions less sensitive to canopy biophysical parameters for leaf chlorophyll estimation using properties optique spectrales des feuilles (PROSPECT) and scattering by arbitrarily inclined leaves (SAILH) simulated data. *Int J Remote Sens.* 2008;29(8):2241-55. <https://doi.org/10.1080/01431160701395328>.
- Chrysafis I, Korakis G, Kyriazopoulos AP, Mallinis G. Retrieval of leaf area index using sentinel-2 imagery in a mixed Mediterranean forest area. *ISPRS Int J Geo-Inf.* 2020;9(11):622. <https://doi.org/10.3390/ijgi9110622>.
- Confalonieri R, Foi M, Casa R, Aquaro S, Tona E, Peterle M, et al. Development of an app for estimating leaf area index using a smartphone. Trueness and precision determination and comparison with other indirect methods. *Comput Electron Agric.* 2013;96:67-74. <https://doi.org/10.1016/j.compag.2013.04.019>.
- Daoud JI. Multicollinearity and regression analysis. Paper presented at: 4th International Conference on Mathematical Applications in Engineering; 2017 Aug 8-9; Kuala Lumpur, Malaysia. Kuala Lumpur: International Islamic University Malaysia, 2017.
- Darvishzadeh R, Wang T, Skidmore A, Vrieling A, O'Connor B, Gara TW, et al. Analysis of sentinel-2 and rapideye for retrieval of leaf area index in a saltmarsh using a radiative transfer model. *Remote Sens.* 2019;11(6):671. <https://doi.org/10.3390/rs11060671>.
- Djamai N, Fernandes R, Weiss M, McNairn H, Goita K. Validation and Comparison of Cropland Leaf Area Index Retrievals from Sentinel-2/MSI Data Using S12P Processor and Vegetation Indices Models. Paper presented at: IGARSS 2019 - 2019 IEEE International Geoscience and Remote Sensing Symposium; 2019 Jul 28-Aug 2; Yokohama, Japan. Yokohama: IEEE, 2019a. pp. 4595-98.
- Djamai N, Zhong D, Fernandes R, Zhou F. Evaluation of vegetation biophysical variables time series derived from synthetic sentinel-2 images. *Remote Sens.* 2019b;11(13):1547. <https://doi.org/10.3390/rs11131547>.
- Durbha SS, King RL, Younan NH. Support vector machines regression for retrieval of leaf area index from multiangle imaging spectroradiometer. *Remote Sens Environ.* 2007;107(1-2):348-61. <https://doi.org/10.1016/j.rse.2006.09.031>.
- Eisfelder C. Modelling net primary productivity and above-ground biomass for mapping of spatial biomass distribution in Kazakhstan [Dissertation]. Oberpfaffenhofen: German Remote Sensing Data Center (DFD) of the German Aerospace Center (DLR); 2013.
- Eisfelder C, Klein I, Bekkulyeva A, Kuenzer C, Buchroithner MF, Dech S. Above-ground biomass estimation based on NPP time-series – A novel approach for biomass estimation in semi-arid Kazakhstan. *Ecol Indic.* 2017;72:13-22. <https://doi.org/10.1016/j.ecolind.2016.07.042>.
- Fang H, Li W, Wei S, Jiang C. Seasonal variation of leaf area index (LAI) over paddy rice fields in NE China: Intercomparison of destructive sampling, LAI-2200, digital hemispherical photography (DHP), and AccuPAR methods. *Agric For Meteorology.* 2014;198-199:126-41. <https://doi.org/10.1016/j.agrformet.2014.08.005>.
- Ghebremicael ST, Smith CW, Ahmed FB. Estimating the leaf area index (LAI) of black wattle from Landsat ETM+ satellite imagery. *South Afr For J.* 2004;201(1):3-12. <https://doi.org/10.1080/20702620.2004.10431769>.
- Jacquemoud S, Bacour C, Poilvé H, Frangi JP. Comparison of four radiative transfer models to simulate plant canopies reflectance: direct and inverse mode. *Remote Sens Environ.* 2000;74(3):471-81. [https://doi.org/10.1016/S0034-4257\(00\)00139-5](https://doi.org/10.1016/S0034-4257(00)00139-5).
- Jin M, Zhang DL. Observed variations of leaf area index and its relationship with surface temperatures during warm seasons. *Meteorol Atmos Phys.* 2002;80:117-29. <https://doi.org/10.1007/s007030200019>.
- Jonckheere I, Fleck S, Nackaerts K, Muys B, Coppin P, Weiss M, et al. Review of methods for in situ leaf area index determination: Part I. Theories, sensors and hemispherical photography. *Agric For Meteorology.* 2004;121(1-2):19-35. <https://doi.org/10.1016/j.agrformet.2003.08.027>.

- Joseph M, Wang L, Wang F. Using landsat imagery and census data for urban population density modeling in Port-au-Prince, Haiti. *GISci & Remote Sens.* 2012;49(2):228-50. <https://doi.org/10.2747/1548-1603.49.2.228>.
- Kamenova I, Dimitrov P. Evaluation of Sentinel-2 vegetation indices for prediction of LAI, fAPAR and fCover of winter wheat in Bulgaria. *Eur J Remote Sens.* 2020;54(sup1):89-108. <https://doi.org/10.1080/2797254.2020.1839359>.
- Kganyago M, Mhangara P, Alexandridis T, Laneve G, Ovakoglou G, Mashiyi N. 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 Sens Lett.* 2020;11(10):883-92. <https://doi.org/10.1080/2150704X.2020.1767823>.
- Kim JH. Multicollinearity and misleading statistical results. *Korean J Anesthesiol.* 2019;72(6):558-69. <https://doi.org/10.4097/kja.19087>.
- Li W, Weiss M, Garric B, Champolivier L, Jiang J, Wu W, et al. Mapping crop leaf area index and canopy chlorophyll content using UAV multispectral imagery: impacts of illuminations and distribution of input variables. *Remote Sens.* 2023;15(6):1539. <https://doi.org/10.3390/rs15061539>.
- Lillesand TM, Kiefer RW, Chipman JW. *Remote sensing and image interpretation.* 7th ed. Hoboken, NJ: Wiley; 2015.
- Ma Y, Zhang Q, Yi X, Ma L, Zhang L, Huang C, et al. Estimation of cotton leaf area index (LAI) based on spectral transformation and vegetation index. *Remote Sens.* 2022;14(1):136. <https://doi.org/10.3390/rs14010136>.
- Magney TS, Eitel JUH, Vierling LA. Mapping wheat nitrogen uptake from RapidEye vegetation indices. *Precision Agric.* 2017;18:429-51. <https://doi.org/10.1007/s11119-016-9463-8>.
- Main-Knorn M, Pflug B, Debaecker V, Louis J. Calibration and validation plan for the 12A processor and products of the Sentinel-2 mission. *Int Arch Photogramm Remote Sens Spatial Inf Sci.* 2015;XL-7/W3:1249-55. <https://doi.org/10.5194/isprsarchives-XL-7-W3-1249-2015>.
- Middinti S, Thumaty KC, Gopalakrishnan R, Jha CS, Thatiparthi BR. Estimating the leaf area index in Indian tropical forests using Landsat-8 OLI data. *Int J Remote Sens.* 2017;38(23):6769-89. <https://doi.org/10.1080/01431161.2017.1363436>.
- Mourad R, Jaafar H, Anderson M, Gao F. Assessment of leaf area index models using harmonized landsat and Sentinel-2 surface reflectance data over a semi-arid irrigated landscape. *Remote Sens.* 2020;12(19):3121. <https://doi.org/10.3390/rs12193121>.
- Muhe S, Argaw M. Estimation of above-ground biomass in tropical afro-montane forest using Sentinel-2 derived indices. *Environ Syst Res.* 2022;11:5. <https://doi.org/10.1186/s40068-022-00250-y>.
- Mwangi P, Karanja F, Kamau P. Analysis of the relationship between land surface temperature and vegetation and built-up indices in Upper-Hill, Nairobi. *J Geosci Environ Prot.* 2018;6:1-16. <https://doi.org/10.4236/gep.2018.61001>.
- Myneni RB, Ramakrishna R, Nemani R, Running SW. Estimation of global leaf area index and absorbed par using radiative transfer models. *IEEE Transactions Geosci Remote Sens.* 1997;35(6):1380-93. <https://doi.org/10.1109/36.649788>.
- O'Hagan A. Bayesian analysis of computer code outputs: a tutorial. *Reliab Eng Syst Saf.* 2006;91(10-11):1290-300. <https://doi.org/10.1016/j.res.2005.11.025>.
- Pasqualotto N, Delegido J, Van Wittenberghe S, Rinaldi M, Moreno J. Multi-crop green LAI estimation with a new simple Sentinel-2 LAI index (SeLI). *Sens.* 2019a;19(4):904. <https://doi.org/10.3390/s19040904>.
- Pasqualotto N, D'Urso G, Bolognesi SF, Belfiore OR, Van Wittenberghe S, Delegido J, et al. Retrieval of evapotranspiration from Sentinel-2: comparison of vegetation indices, semi-empirical models and SNAP biophysical processor approach. *Agronomy.* 2019b;9(10):663. <https://doi.org/10.3390/agronomy9100663>.
- Pizaña JMG, Hernández JMN, Romero NC. Remote sensing-based biomass estimation. In: Marghany M, editor. *Environmental applications of remote sensing.* InTechOpen; 2016. p. 3-40.
- Reygadas Y, Jensen JLR, Moisen GG, Currit N, Chow ET. Assessing the relationship between vegetation greenness and surface temperature through granger causality and impulse-response coefficients: a case study in Mexico. *Int J Remote Sens.* 2020;41(10):3761-83. <https://doi.org/10.1080/01431161.2019.1711241>.
- Richter K, Atzberger C, Vuolo F, D'Urso G. Evaluation of Sentinel-2 spectral sampling for radiative transfer model based LAI estimation of wheat, sugar beet, and maize. *IEEE J Sel Top Appl Earth Observat Remote Sens.* 2011;4(2):458-64. <https://doi.org/10.1109/JSTARS.2010.2091492>.
- Sikand MV, Stamnes K. Radiative transfer models: available tools, patents and commercial potential to remote sensing. *Recent Pat Sp Technol.* 2012;2(1):49-56. <https://doi.org/10.2174/1877611611202010049>.
- Singh J, Singh L, Kumar A. Estimation of leaf area by mobile application: fast and accurate method. *Pharma Innovation J.* 2021;10(4S):272-5. <https://doi.org/10.22271/tpi.2021.v10.i4Se.6066>.
- Sishodia RP, Ray RL, Singh SK. Applications of remote sensing in precision agriculture: a review. *Remote Sens.* 2020;12(19):3136. <https://doi.org/10.3390/rs12193136>.
- Sola I, García-Martín A, Sardonis-Pozo L, Álvarez-Mozos J, Pérez-Cabello F, González-Audicana M, et al. Assessment of atmospheric correction methods for Sentinel-2 images in Mediterranean landscapes. *Int J Appl Earth Obs Geoinf.* 2018;73:63-76. <https://doi.org/10.1016/j.jag.2018.05.020>.
- Stenberg P, Linder S, Smolander H, Flower-Ellis J. Performance of the LAI-2000 plant canopy analyzer in estimating leaf area index of some Scots pine stands. *Tree Physiol.* 1994;14(7-9):981-95. <https://doi.org/10.1093/treephys/14.7-8-9.981>.
- Upreti D, Huang W, Kong W, Pascucci S, Pignatti S, Zhou X, et al. A comparison of hybrid machine learning algorithms for the retrieval of wheat biophysical variables from Sentinel-2. *Remote Sens.* 2019;11(5):481. <https://doi.org/10.3390/rs11050481>.
- Vafaei S, Fathizadeh O, Puletti N, Fadaei H, Baqer Rasooli S, Vaglio Laurin G. Estimation of forest leaf area index using satellite multi-spectral and synthetic aperture radar data in Iran. *iForest.* 2021;14:278-84. <https://doi.org/10.3832/ifer3633-014>.
- Verrelst J, Rivera JP, Veroustraete F, Muñoz-Marí J, Clevers JGPW, Camps-Valls G, et al. Experimental Sentinel-2 LAI estimation using

- parametric, non-parametric and physical retrieval methods – a comparison. *ISPRS J Photogrammetry Remote Sens.* 2015;108:260-72. <https://doi.org/10.1016/j.isprsjprs.2015.04.013>.
- Verrelst J, Sabater N, Rivera JP, Muñoz-Marí J, Vicent J, Camps-Valls G, et al. Emulation of leaf, canopy and atmosphere radiative transfer models for fast global sensitivity analysis. *Remote Sens.* 2016;8(8):673. <https://doi.org/10.3390/rs8080673>.
- Weiss M, Baret F. S2ToolBox Level 2 products: LAI, FAPAR, FCOVER - version 1.1. S2 Toolbox Level 2 Prod algorithm. 2016;V1.1:1-53.
- Xie Q, Dash J, Huang W, Peng D, Qin Q, Mortimer H, et al. Vegetation indices combining the red and red-edge spectral information for leaf area index retrieval. *IEEE J Sel Top Appl Earth Observat Remote Sens.* 2018;11(5):1482-93. <https://doi.org/10.1109/JSTARS.2018.2813281>.
- Xie Q, Dash J, Huete A, Jiang A, Yin G, Ding Y, et al. Retrieval of crop biophysical parameters from Sentinel-2 remote sensing imagery. *Int J Appl Earth Observat Geoinf.* 2019;80:187-95. <https://doi.org/10.1016/j.jag.2019.04.019>.
- Xu J, Quackenbush LJ, Volk TA, Im J. Forest and crop leaf area index estimation using remote sensing: research trends and future directions. *Remote Sens.* 2020;12(18):2934. <https://doi.org/10.3390/rs12182934>.
- Yao X, Wang N, Liu Y, Cheng T, Tian Y, Chen Q, et al. Estimation of wheat LAI at middle to high levels using unmanned aerial vehicle narrowband multispectral imagery. *Remote Sens.* 2017;9(12):1304. <https://doi.org/10.3390/rs9121304>.
- Zhang J, Qiu X, Wu Y, Zhu Y, Cao Q, Liu X, et al. Combining texture, color, and vegetation indices from fixed-wing UAS imagery to estimate wheat growth parameters using multivariate regression methods. *Comput Electron Agric.* 2021;185:106138. <https://doi.org/10.1016/j.compag.2021.106138>.
- Zhang Y, Yang J, Liu X, Du L, Shi S, Sun J, et al. Estimation of multi-species leaf area index based on Chinese GF-1 satellite data using look-up table and gaussian process regression methods. *Sens.* 2020;20(9):2460. <https://doi.org/10.3390/s20092460>.
- Zheng G, Moskal LM. Retrieving leaf area index (LAI) using remote sensing: theories, methods and sensors. *Sens.* 2009;9(4):2719-45. <https://doi.org/10.3390/s90402719>.