

**Office of Satellite and Product Operations
Environmental Satellite Processing Center**



**VIIRS Leaf Area Index Algorithm Theoretical
Basis Document**

Version 1.1

April 21, 2025

**U.S. Department of Commerce
National Oceanic and Atmospheric Administration
National Environmental Satellite, Data, and Information Service
Office of Satellite and Product Operations**

Authors

Heshun Wang (UMD/ESSIC, College Park, MD)

Yunyue Yu (NOAA/NESDIS/STAR, College Park, MD)

Yuan Zhou (UMD/ESSIC, College Park, MD)

Ivan Csiszar (NOAA/NESDIS/STAR, College Park, MD)

Changes/Revisions Record

This algorithm theoretical basis document is changed as required to reflect system, operational, or organizational changes. Modifications made to this document are recorded in the Changes/Revisions Record below. This record will be maintained throughout the life of the document.

Version Number	Date	Description of Change/Revision	Section/Pages Affected	Changes Made by Name/Title/Organization
1.0	01/31/2024	New document created		
1.1	04/21/2025	Cover page, Authors, Changes/Revisions Record, and Preface added to Version 1.0. Updated Table of Contents, List of Tables, List of Figures, headers, footers, and margins.	All	Hannah Bowie, Technical Writer, ERT
1.1	04/25/2025	Quality Assurance	All	Clint Sherwood, Quality Assurance Manager, ERT

Preface

This document comprises the National Oceanic and Atmospheric Administration (NOAA) National Environmental Satellite, Data, and Information Service (NESDIS), Office of Satellite and Product Operations (OSPO), publication of this VIIRS Leaf Area Index (LAI) Algorithm Theoretical Basis Document. This document reflects current operations for the DOC/NOAA/NESDIS Environmental Satellite Processing Center (ESPC) (NOAA5045) information technology systems. This document describes the established ESPC procedure for LAI users in accordance with Federal, DOC, NOAA, NESDIS and OSPO requirements.

The published version of this document can be found at the OSPO SharePoint Products Library.

Table of Contents

INTRODUCTION	1
Product Overview	1
1.1.1 Product Description	1
1.1.2 Product Requirements	2
Satellite Instrument Description	2
ALGORITHM DESCRIPTION.....	3
Processing Outline	3
Theoretical Description.....	5
2.2.1 Introduction.....	5
2.2.2 Machine learning algorithms	6
2.2.3 Data preparation.....	11
2.2.4 LAI retrieval and compositing	14
2.2.5 Temporal smoothing and gap filling.....	15
Algorithm Input	16
Algorithm Output.....	18
Performance Estimates.....	19
2.5.1 in-situ LAI validation.....	19
2.5.2 Inter-comparison with existing products	22
2.5.3 LAI performance in Noah-MP model.....	23
ASSUMPTIONS AND LIMITATIONS	24
Performance Assumptions	24
Potential Improvements	24
REFERENCES	26

List of Tables

Table 1 - JPSS LAI Product Requirements.....	2
Table 2 - VIIRS Surface Reflective bands and configurations.....	3
Table 3 - SR compositing selection priority	12
Table 4 - New biome type conversion scheme from IGBP classes	12
Table 5 - List of dynamic input data sets used in the VIIRS LAI algorithm.....	16
Table 6 - List of static input used in the VIIRS LAI algorithm.....	17
Table 7 - VIIRS SR output science datasets	18

List of Figures

Figure 1 - Leaf Area Index processing architecture.....	5
Figure 2 - VIIRS spectral response (I1, I2 and I3) and typical vegetation reflectance.....	8
Figure 3 - Red/NIR SR scattering plot and the convex hull	9
Figure 4 - Random Forest model performance for each biome type	10
Figure 5 - The global biome type map used in LAI retrieval	14
Figure 6 - LAI weekly processes based on the daily retrieval.....	15
Figure 7 - Surface Reflectance validation results at AERONET.....	17
Figure 8 - The ground LAI upscaling using the high-resolution LAI (Left: 30m Landsat LAI, Right: aggregated 300m LAI) as bridge	20
Figure 9 - The LAI product preliminary validation at NEON sites.....	21
Figure 10 - Preliminary inter-comparison with VNP15 LAI at the BELMANIP2 sites	23

INTRODUCTION

This document outlines the scientific foundation, design, and expected performance of the Leaf Area Index (LAI) algorithm for the in-operation Suomi National Polar-orbiting Partnership (S-NPP), NOAA-20, NOAA-21 and the future satellites in Joint Polar Satellite System (JPSS).

It includes a concise overview of the product's background, requirements, and the associated satellite instrument in the introductory section. The second section provides a comprehensive overview of the LAI algorithm, including its operational processing flow. This section details the theoretical framework and processing methods, input and output data for the algorithm, as well as plans for evaluation, validation, and anticipated performance. Section 3 delves into the assumptions and limitations inherent to the algorithm, also discussing potential enhancements. The final section enumerates the references utilized.

Product Overview

1.1.1 Product Description

LAI is defined as one half the total green leaf area per unit horizontal ground surface area. It is an essential climate variable driving water and carbon fluxes, and energy exchanges, playing an important role in the models of the climate, hydrology, ecology, et.al. As a fundamental attribute of vegetation, LAI has been listed as an essential climate variable by the global climate change research community (GCOS, 2011).

As the operational land surface model (LSM) in the Global Forecast System (GFS), the Noah LSM takes the green vegetation fraction (GVF) as input for the vegetation dynamic characterization while LAI simply used some constant value for each biome; however, in the next generation LSM, Noah with multi-parameterization (Noah-MP), LAI becomes a key vegetation parameter, and high-quality satellite LAI product has been proved with a significant impact on Noah-MP performance, in this context, a new JPSS LAI operational product is under development to support the LSM and other earth science research and applications.

This document was prepared by the Center for Satellite Applications and Research (STAR) Land Product Development team led by Dr. Yunyue Yu and in consultation with primary data users, e.g. NOAA Environmental Modeling Center. The responsible entity for bookkeeping, accessibility and distribution of this document is the Operational Products Development Branch (OPDB) of the NESDIS STAR Satellite Meteorology and Climatology Division (SMCD). The intended users of this document are project managers, product users, requirement reviewers and code reviewers.

1.1.2 Product Requirements

Product requirements initially proposed by the land surface model users and Table 1 provides the LAI product requirements for the JPSS mission.

Table 1 - JPSS LAI Product Requirements

Attribute	Threshold	Objective
Geographic coverage	Clear sky condition, land surface	All weather condition, land surface
Vertical Coverage	NA	NA
Refresh rate	8-day	Daily rolling 8-day
Horizontal Cell Size	1 km	500 m
Mapping Uncertainty	1 km	500 m
Measurement Range	0-10	0-10
Accuracy	15%	10%
Precision	18%	13%
Uncertainty*	20%	15%

*According to the world meteorological organization (WMO, <https://gcos.wmo.int/>), the uncertainty requirement for global numerical weather prediction (NWP) and High-Resolution NWP application, the threshold is 20%, breakthrough is 10%, goal is 5%.

According to GCOS-200 (Plan 2016), required measurement uncertainty is: maximum (15%).

Satellite Instrument Description

The LAI algorithm primarily uses the data from the VIIRS (Visible Infrared Imaging Radiometer Suite) instrument on the Suomi National Polar-orbiting Partnership (S-NPP) platform and on subsequent satellites of the Joint Polar Satellite System (JPSS).

S-NPP was launched on October 28, 2011, while NOAA-20 and NOAA-21 were launched on November 18, 2017, and November 10, 2022, respectively. These satellites are in a sun-synchronous orbit with a 1:30 pm ascending-node orbit, spaced half an orbit apart, at altitudes of approximately 830 km. VIIRS will also be flown on the JPSS-3 and -4 satellites in the future.

The VIIRS instrument aboard these satellites is a whiskbroom scanning radiometer. It features a swath width of 3060 km, enabling full daily coverage of the Earth's surface on both the day and night sides. VIIRS is equipped with 22 spectral bands that span the spectrum from 0.41 μm to 12.5 μm . This range includes 16 moderate-resolution bands (M-bands) with a spatial resolution of 750 m at nadir, 5 imaging-resolution bands (I-bands) with a spatial resolution of 375 m at nadir, and one panchromatic Day/Night Band (DNB) with a 750 m spatial resolution throughout the scan. The surface reflective bands

information could be found in Table 2, in LAI retrieval, only the image bands (I1, I2 and I3) are used to generate LAI with original spatial resolution of 500 m.

The Level-1b Sensor Data Records (SDR) from VIIRS are the calibrated and geolocated radiance and reflectance data produced from the Raw Data Records. The geolocation data will be used in this product. Based on the SDR, many Level-2 Environmental Data Records (EDR) are generated, the surface reflectance EDR is the primary input for the LAI product, and the annual surface type EDR will be used for the static biome type derivation.

Table 2 - VIIRS Surface Reflective bands and configurations

VIIRS Band	wavelength (μm)	Bandwidth (μm)	SNR*	Spatial resolution (m)
M1	0.412	0.402-0.422	352/316	750m
M2	0.445	0.436-0.454	380/409	
M3	0.488	0.478-0.488	416/414	
M4	0.555	0.545-0.565	362/315	
M5	0.672	0.662-0.682	242/360	
M7	0.865	0.846 - 0.885	215/340	
M8	1.240	1.23 - 1.25	74	
M10	1.61	1.58 - 1.64	83	
M11	2.25	2.23 - 2.28	10	
I1	0.64	0.6 - 0.68	119	
I2	0.865	0.85 - 0.88	150	
I3	1.61	1.58 - 1.64	6	

*M1, M2, M3, M5 and M7 are dual gain with high gain value followed by the low gain Signal-to-noise ratio (SNR), the rest bands are single gain. SNR data are for SNPP and should be slight difference for NOAA-20 and NOAA-21.

ALGORITHM DESCRIPTION

Processing Outline

Satellite LAI datasets, recorded over the past two decades, have been utilized extensively across various applications. Leveraging the legacy of established satellite products like the Moderate Resolution Imaging Spectroradiometer (MODIS), Global Land Surface Satellite (GLASS), and Geoland2/BioPar (GEOV2) LAI products, a data-driven

methodology has been developed to obtain near-real-time LAI from VIIRS observations. Prior to implementation, a machine learning algorithm is tuned and trained based on a comprehensive suite of representative datasets.

The VIIRS LAI product is designed to be a temporally smoothed, global, gap-free dataset. The operational procedure is segmented into three phases, as depicted in Figure 1:

(1) Daily Surface Reflectance Generation: Utilizing the VIIRS gridding tool, granule data is mapped onto a global grid in a sinusoidal projection. The surface reflectance (SR) compositing process then identifies and selects the highest quality SR and corresponding angles for each grid cell.

(2) Daily LAI Retrieval: A previously trained machine learning algorithm performs the clear-sky LAI retrieval, leveraging the daily SR together with auxiliary data.

The first two steps are daily processing, with up to 8 days' data being sustained for the weekly processing, which will be run every 8 days.

(3) 8-Day LAI Compositing and Post-Processing: From the daily LAI outputs, the optimal quality LAI is chosen for each 8-day interval. Subsequently, a temporal smoothing and gap-filling (TSGF) procedure is applied to produce the final, gap-free product.

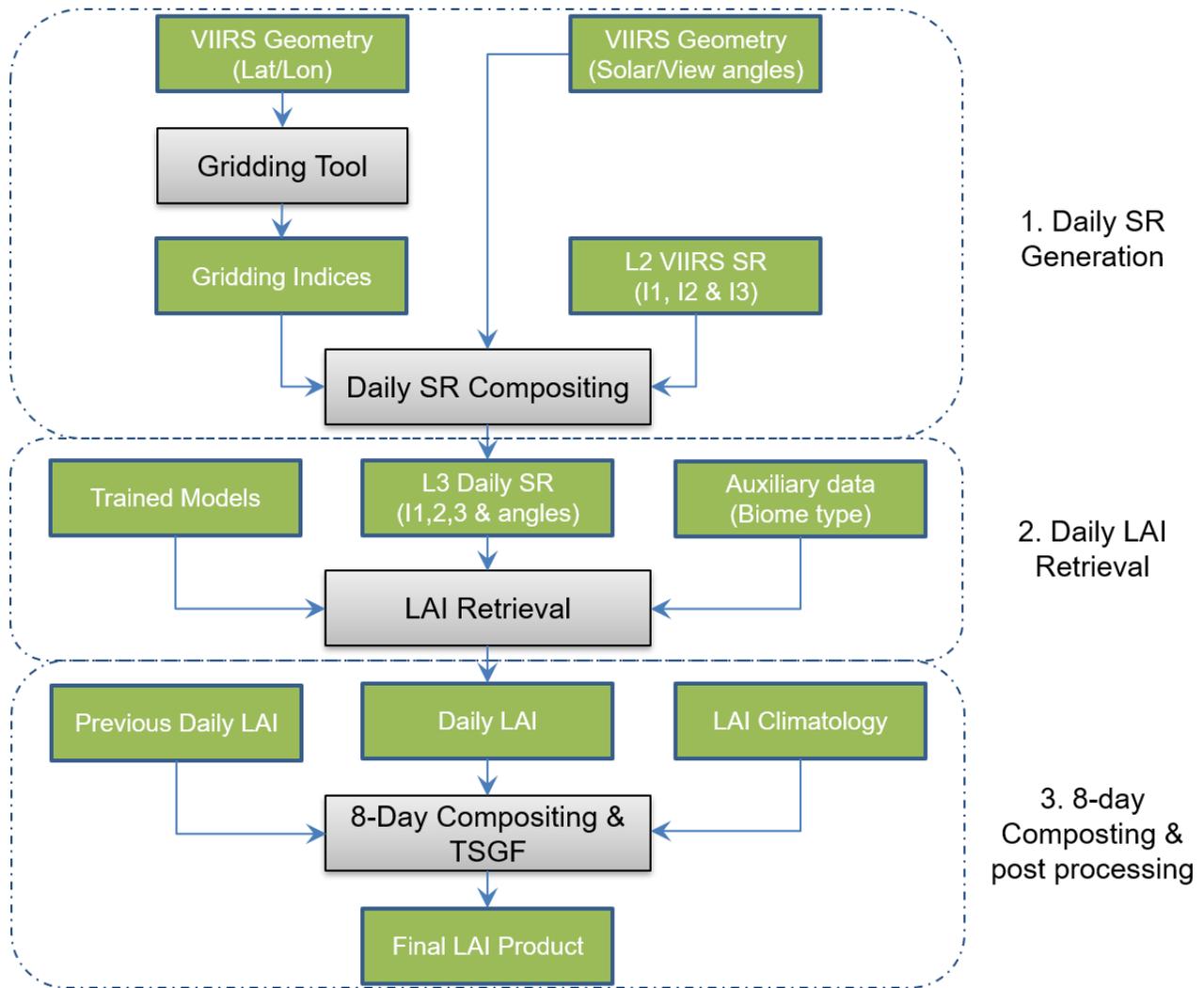


Figure 1 - Leaf Area Index processing architecture

Theoretical Description

2.2.1 Introduction

Supervised machine learning algorithms represent a data-driven approach that is increasingly being applied in the retrieval of the land or atmospheric parameters from satellite imagery. The performance of these algorithms highly depends on the quality (both the accuracy and representativeness) of the training datasets. Over the years, several LAI products have been developed, offering a substantial advantage for the machine learning algorithms. Nonetheless, these products exhibit significant variations in both spatial distribution and temporal series, with each product having its shortcomings. And the limited ground measurements further restrict the use of a single product in applications. Exploring the best LAI estimation from multiple products and developing a

data-driven algorithm trained on fused high-quality LAI data and VIIRS surface reflectance emerges as a promising strategy to produce near-real-time VIIRS products.

2.2.2 Machine learning algorithms

To select an effective machine learning algorithm for VIIRS LAI retrieval, a comprehensive collection of relevant datasets has been prepared for training purposes. Given that LAI is intended to be an operational product where both efficiency and accuracy are important, the evaluation of the models and coefficients will carefully balance the trade-off between uncertainty and operational efficiency.

Tree-based algorithms have proven exceptionally adept at extracting land surface parameters from remote sensing data. Among these, the Random Forest algorithm stands out for its computational efficiency, making it particularly suitable for processing the extensive datasets generated by global observations. This efficiency, coupled with its effectiveness, positions the Random Forest algorithm as the recommended choice for estimating the VIIRS LAI.

The Random Forest model, a supervised learning algorithm, employs ensemble learning for both classification and regression tasks. It builds numerous decision trees during training and averages their predictions to yield robust outcomes. Each tree in the ensemble is constructed from a unique, randomly selected subsample of the training set and utilizes a random subset of input features. This inherent randomness not only mitigates overfitting, thus enhancing generalization, but also bolsters the model's resilience against noisy data.

Training datasets for LAI Estimation Using Machine Learning

(1) NASA MODIS/VIIRS LAI: Derived from daily MODIS red and near-infrared (NIR) surface reflectance via biome-specific Look-Up Tables (LUTs) from a three-dimensional radiative transfer model. A backup NDVI empirical method is used if the main algorithm fails. Despite quality control and temporal compositing for an 8-day, 500m resolution LAI product, some cloud-contaminated pixels remain, leading to outliers and fluctuations over time.

(2) Copernicus Global Land Service GEO v2 LAI: This enhanced version employs an Artificial Neural Network (ANN) trained on PROBA-V red, NIR, and Short-Wave Infrared (SWIR, 1.6 μ m) surface reflectance. The training integrates a weighted average of CYCLOPES (V3.1) and MODIS LAI (Collection 5) data. The product undergoes filtering, temporal compositing, and gap filling for a gap-free, smoothed time series.

To be noted, the GEOv2 training LAI taking MODIS LAI for dense vegetation ($LAI > 4$), the weighting decreases progressively for LAI values below 4, when CYCLOPES LAI is estimated from model training data generated by SAIL model. GEOv2 LAI is produced as a decadal product (10 days) at 1/112 degree, which is different from the proposed spatial and temporal resolution, so a linear interpolation method is used to convert to 8-day

refresh rate data, and a nearest neighbor method is used to convert the spatial resolution to 500m sinusoidal grid.

(3) GLASS LAI (Version 6): Different from MODIS and GEO v2, which utilize daily data for LAI derivation, GLASS LAI employs a deep learning approach using a Bidirectional Long Short-Term Memory (BiLSTM) network using long-term MODIS surface reflectance data as input. The training model incorporates a fused LAI time-series from MODIS, GEO v1, and GLASS LAI (version 5). As a result, the final GLASS LAI products are provided gap-free at both 250m and 500m resolutions. Due to the smoothing process applied to the trained LAI data, the output LAI is also expected to exhibit a smoothed quality.

To be noted, the expected LAI product is the true LAI as the physical definition described, however, some LAI product base on gap fraction theory is with the concept of effective LAI, such as CYCLOPES used in GEOv2 and GLASS v5, which need adjustment using clumping index to match the true LAI such as MODIS. Therefore, all these three LAI datasets are defined as true LAI and ensure this VIIRS LAI is true value. The retrieved LAI from optical remote sensing observations is mainly corresponding to the green element: the correct term to be used would be GAI (Green Area Index) although we propose to still use LAI for the sake of simplicity.

The primary estimator for predicting LAI is the VIIRS surface reflectance measurements. For LAI estimation, not only the traditional red (VIIRS I1) and near-infrared (VIIRS I2) bands are utilized, but the SWIR band (VIIRS I3) is also incorporated. The SWIR band located at 1.6 μ m (as Figure 2 shows) is known for its sensitivity to leaf water content, which many researchers have found beneficial for LAI derivation. Given that reflectance characteristics are highly anisotropic, and LAI is inherently a structural parameter, both solar and viewing angles are included for the Bidirectional Reflectance Distribution Function (BRDF) adjustment process. Supported by a feature importance evaluation, the vegetation indices like the NDVI and the NDWI (Normalized Difference Water Index) are integrated into the model as well to enhance performance, owing to their proven strong correlations with LAI.

LAI is strongly nonlinearly related to the reflectance, with the relationship being dependent on biome type. Different biomes exhibit unique vertical structures, soil types, and clumping indices, making the prior information on biome types crucial for generating a global LAI product. After conducting a clustering analysis, 6 biome types are proposed in the algorithm to train the model separately, and the biome data will be used in daily retrieval to determine the model for each type. The details about the biome data will be introduced in 2.2.3.

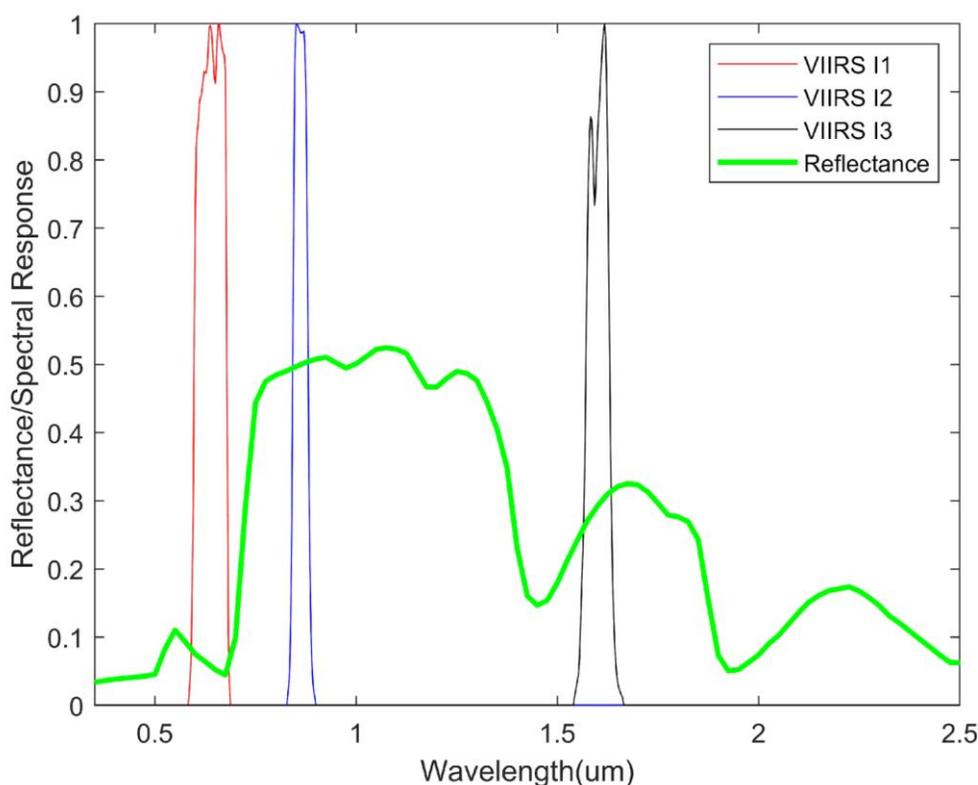


Figure 2 - VIIRS spectral response (I1, I2 and I3) and typical vegetation reflectance

For the model training, a set of globally distributed datasets spanning 2 years were used to encompass at least two growing seasons. This was done to ensure the collection of sufficient data that is both robust and representative. In addition to utilizing data from the 445 Benchmark Land Multisite Analysis and Intercomparison of Products (BELMAINIP version 2) sites, a global sampling method was employed to select training data from across the world. Selection criteria included ensuring the pixel was in a land area (excluding coastal or permanently snow-covered areas), that the surrounding biome type and LAI value were homogeneous, and that the distribution was uniform across the global area.

To enhance the efficiency of the training data, several processes were implemented:

(1) Data Screening: The quality of the training data is crucial for model performance. Several screening methods were applied, including the use of quality flags with LAI and SR (Surface Reflectance) data. For MODIS LAI data, only data derived from the main algorithm under clear sky conditions were used. Both unsaturated and potentially saturated LAI data were utilized. GEOv2 and GLASS data, being postprocessed, were screened to select the most valid data for fusion. Similarly, VIIRS SR data were selected based on quality flags to ensure reliability. The LAI data fusion method was also used to enhance data quality by comparing three different products. For SR data, the NDVI-LAI

relationship was used to exclude potential outliers by screening out pixels with NDVI values that were too high or too low for the given LAI.

(2) Biome Type Representation: Given the sensitivity of LAI retrieval to biome type, a biome type screening process was also conducted. For each biome, a Red-NIR scattering plot (as Figure 3 shows) was used to define a convex hull (polygon) that encompassed 97% of the data points, with the remaining points deemed less representative of the biome. During LAI retrieval, input data were tested to ascertain whether they fell within the biome's polygon, with the results recorded in the quality flag for diagnostic purposes.

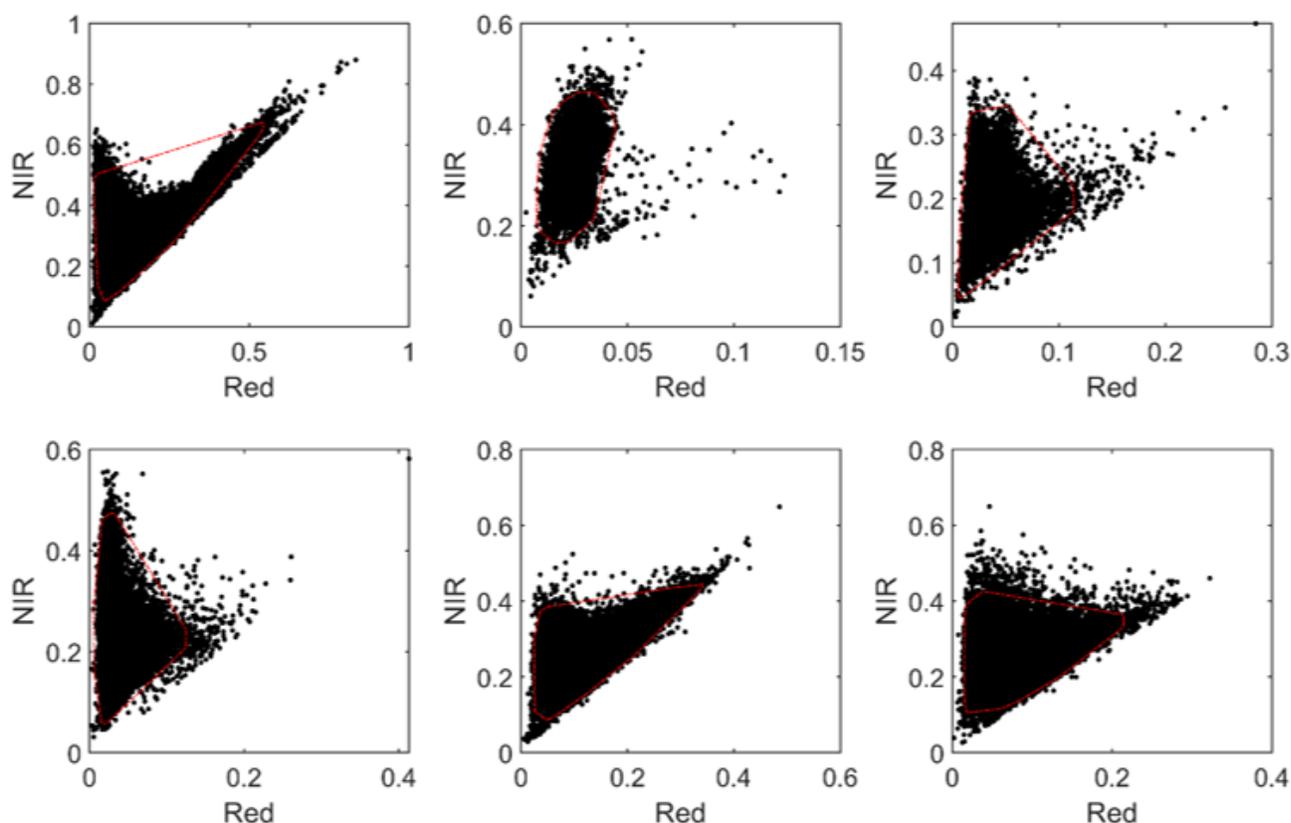


Figure 3 - Red/NIR SR scattering plot and the convex hull

(3) Sample Balancing: After screening, a process of balancing was applied to the high-quality training data to minimize redundancy and enhance representation. Initially, a clustering experiment using the K-means clustering analysis (without supervision) was conducted to reduce data considered duplicated, resulting in over 10,000 clusters. Subsequently, the distribution of biome type, day of the year, LAI value, and locations were adjusted to match a global average distribution. It is noted that areas with high LAI are more affected by the atmosphere than those with low LAI; therefore, retaining more high LAI training data helps to prevent overfitting issues.

We conducted a thorough evaluation of various machine learning algorithms, including Cubist, Random Forest, Gaussian Process Regression, and Artificial Neural Networks (ANN). This assessment involved a rigorous comparison of each model's performance in terms of accuracy, consistency, and computational efficiency. And Random Forest emerged as the most suitable algorithm for our purposes. The test indicated its robustness in capturing the complex nonlinear relationships inherent in environmental data, combining the predictions from multiple decision trees to improve the overall prediction accuracy, despite the algorithm's complexity, it can run relatively quickly, making it feasible for large-scale applications and operational use.

Tuning is the task of finding optimal hyperparameters for a learning algorithm for a considered dataset. In the exploratory analysis, we evaluated the influence of hyperparameters on model performance through 10-fold cross-validation procedures. The Random Forest algorithm consists of multiple decision trees, key variables tested were the number of trees in the forest (ntree), the number of features (mtry) considered for splitting at each node within the trees, and the minimum size of the terminal nodes (leaves) of the trees. Increasing the number of trees generally improves model performance up to a certain point but results in longer training time. And a larger mtry can decrease bias but increase variance, A smaller node size allows for catching more details at the risk of overfitting. So, it's important to find a balance when tuning the parameters.

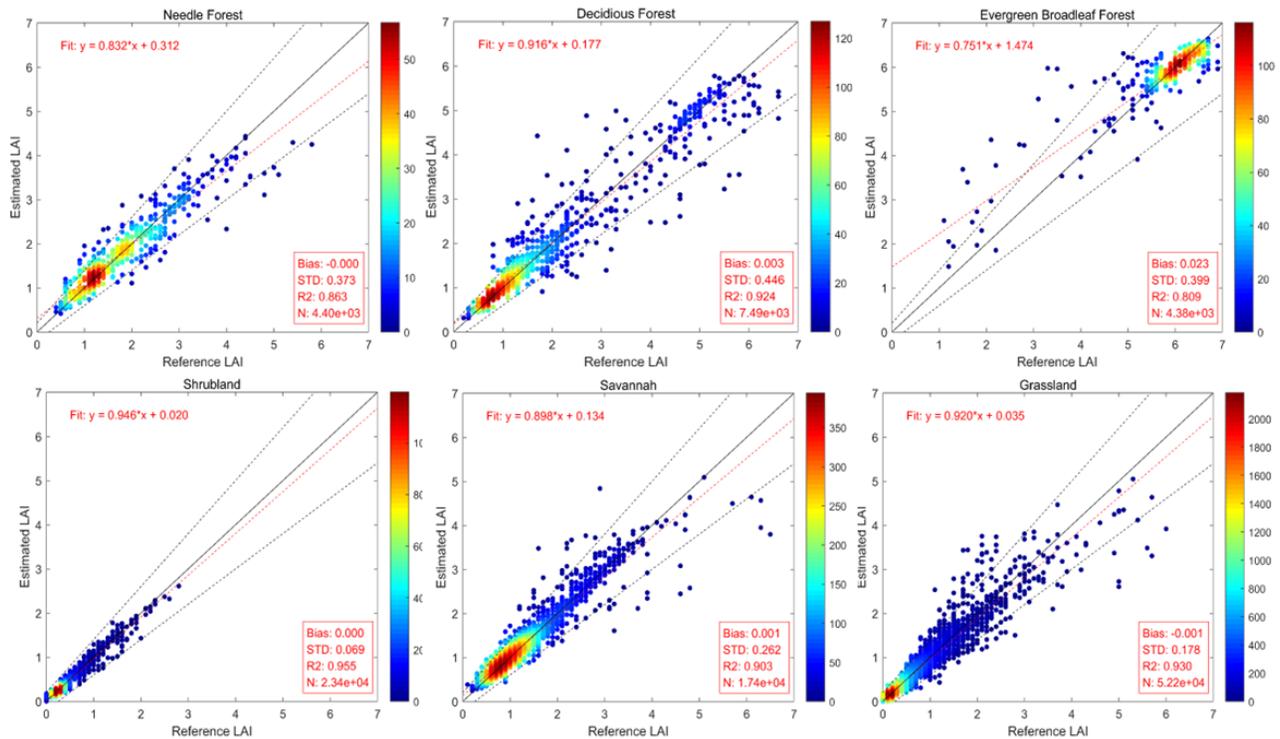


Figure 4 - Random Forest model performance for each biome type

The random forest models were developed to estimate the LAI for six biome types, utilizing two years of global data. The model's parameters were optimized and saved specifically for VIIRS LAI estimation. Validation with independent datasets indicates strong model performance, as illustrated in Figure 4. Notably, forests—particularly evergreen broadleaf forests—exhibited considerable uncertainty due to their complex vertical structure. In these rainforests, the LAI can exceed 5, and the visible/near-infrared bands are prone to saturation. Compared to shrublands and grasslands, savannahs also displayed greater uncertainty, underscoring the model's varied accuracy across different biomes

2.2.3 Data preparation

LAI retrieval is performed based on the daily global gridded surface reflectance, incorporating two key subroutines: mapping VIIRS granule data onto a fixed grid of a specific projection and compositing the SR values within each grid cell. The mapping process utilizes an independent gridding tool, which has been operationally applied to Land Surface Temperature and Albedo Level 3 products. This gridding tool is designed to transform VIIRS granule data into stacked layers within predefined tiles. The global grid comprises 72 by 72 tiles, with each tile representing a 5 by 2.5-degree area. The tool takes latitude and longitude as inputs and forwards maps the granule pixels to a sinusoidal grid at 500m resolution.

To address the bow-tie effect at the swath edges, a gap-filling approach is employed, ensuring accurate mapping. The gridding tool, designed as a generic solution, maintains detailed indices for tile-to-granule and granule-to-tile mappings. These indices facilitate the precise mapping of granule SR (I1, I2 and I3 band) data along with the solar/satellite zenith angles, azimuth angles to the tile grid prior to the compositing process. For detailed methodology of the gridding process, please refer to the Algorithm Theoretical Basis Document For Gridded VIIRS Land Surface Temperature and Albedo Production (https://www.star.nesdis.noaa.gov/jpss/documents/ATBD/ATBD_Gridded_VIIRS_LST_LSA_v1.pdf)

For the daily SR compositing, when multiple observations exist within some grid cells, a selection priority criterion is applied to choose the best quality data. The quality assessment involves a test scoring range from 0 to 10 (see Table 3). The surface reflectance will go through the test 0 to 10, when failing in the test the data receives the score of the test index, and the pixel with the highest score is selected for the daily SR compositing, if the candidates share the same score, the one with lower VZA is preferred. This ensures that the composited SR data represents the highest quality observation available for each grid cell.

Daily gridded SR serves as an intermediate dataset, and will be used to do the daily LAI retrieval. For computing efficiency, the output of daily SR saves in accordance with MODIS Sinusoidal Tile Grid at 500m resolution, there are around 315 tiles with land surface data are stored, each tile (2400*2400 grids) is 10 degrees by 10 degrees at the

equator. In addition to three VIIRS Image Band SR and four solar and view angles, two 8-bit quality flags are recorded as reference for the further processing. This includes the cloud mask, high AOD flag, cloud shadow mask flag et, al.

For the LAI retrieval, some auxiliary static data are needed, one is the biome type, which are derived from multiple years of VIIRS annual surface type (AST), which utilizes the 17-class scheme developed by the International Geosphere Biosphere Program (IGBP). However, the IGBP classification does not directly meet the requirements for LAI retrieval. After conducting a clustering analysis, IGBP types that show similar performance in LAI retrieval are merged into new categories. The Evergreen Broadleaf Forest category is kept as is due to its distinct characteristics, while all low-height, uniformly distributed vegetation types are consolidated into a single category. In utilizing VIIRS AST data, a climatology is generated from all available data. Subsequently, the new biome type data is reassigned according to the plan outlined in Table 4. And the regrouped biome map as shown in Figure 5.

Table 3 - SR compositing selection priority

0	Filled value or out of range [0,1]
1	I1, I2 and I3 SDR QF at least one is marked bad
2	Large View Zenith Angle (≥ 60)
3	Large Solar Zenith Angle (≥ 85)
4	Confidently Cloudy
5	Probably Cloudy
6	Probably Clear
7	Cloud shadow
8	High AOD quantity
9	Not AOD climatology
10	No Snow Present

Since the biome class used in the Noah-MP model are different from the one proposed in this document, evaluation will be conducted to determine if the consistent biome classes are needed in the future.

Table 4 - New biome type conversion scheme from IGBP classes

New Biome types	IGBP classes
Needleleaf Forests	IGBP = 1. Evergreen Needleleaf Forests IGBP = 3. Deciduous Needleleaf Forests
Deciduous Broadleaf Forest	IGBP = 4. Deciduous Broadleaf Forests. IGBP = 5. Mixed Forests
Evergreen Broadleaf Forest	IGBP = 2. Evergreen Broadleaf Forests

shrubland	IGBP = 6. Closed Shrublands. IGBP = 7. Open Shrublands
Savannah	IGBP = 8. Woody Savannas. IGBP = 9. Savannas
Grassland cropland & others	IGBP = 10. Grasslands. IGBP = 11. Permanent Wetlands IGBP = 12. Croplands. IGBP = 13. Urban and Built-Up Lands IGBP = 14. Cropland/Natural Vegetation Mosaics. IGBP = 15. Snow and Ice IGBP = 16. Barren. IGBP = 17. Water;

Another static auxiliary data is the land water mask, which is used to identify all the land grids, excluding all the non-land grids and filling all the gap in the land. The mask is based on the NASA DEM datasets derived from the Shuttle Radar Topography Mission (SRTM) with 15 arc-second resolutions. The ocean coastlines and lake shorelines along with the land are marked as land, all else as water.

The last but very important auxiliary data is the LAI climatology, which is used for temporal smoothness and gap filling. Before the VIIRS LAI dataset was sufficiently populated for climatology purposes, the GLASS LAI dataset from the latest 10 years (2012-2021) was utilized, offering a resolution of 500m and an 8-day refresh rate to create an LAI climatology. To ensure each land grid contained valid LAI data, a comprehensive gap-filling process was implemented. Initially, gaps were filled using valid data from the same biome and latitude within the tile (10° by 10°). If no data were available within this range, the search expanded globally to any location at the same latitude.

Both the land water mask and LAI climatology are saved in the same MODIS sinusoidal tiles, ensuring that the consistent subsets and computing efficiency.

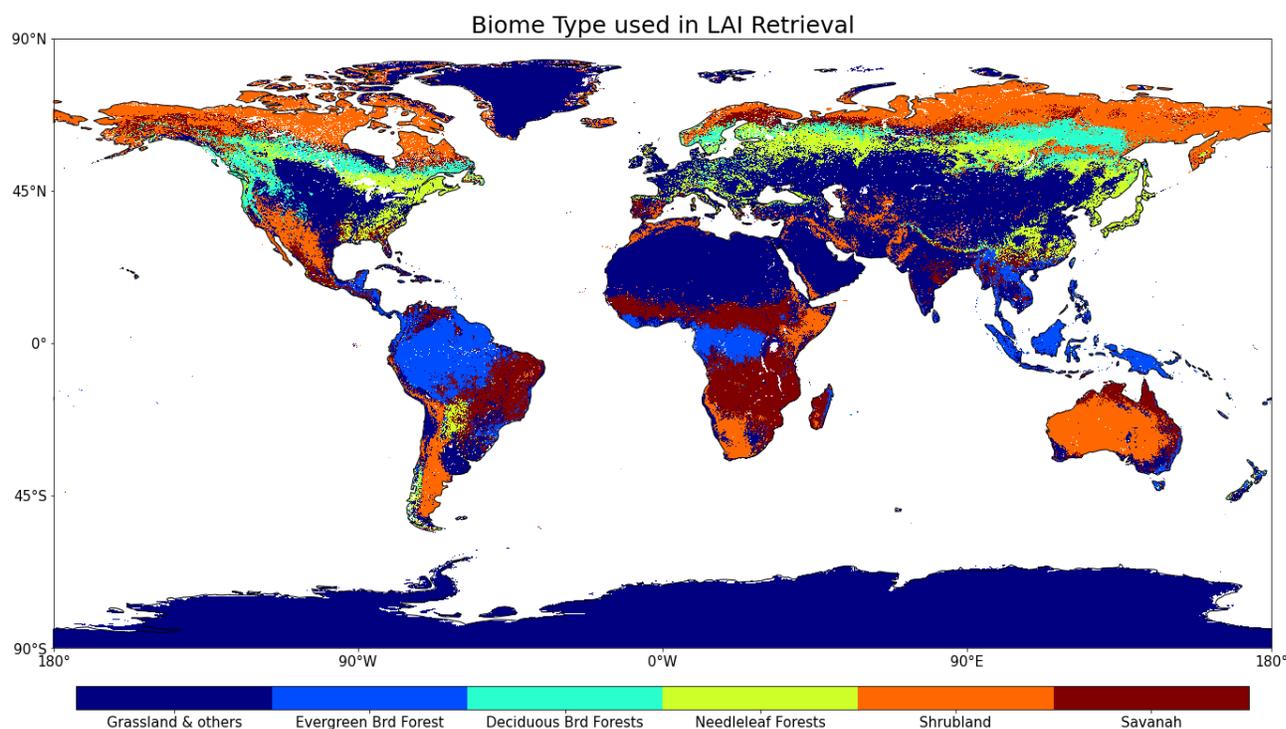


Figure 5 - The global biome type map used in LAI retrieval

2.2.4 LAI retrieval and compositing

The Daily LAI retrieval is the most fundamental process of the product, it directly gets the LAI information from the SR observation, all the post processing is based on this data.

The daily LAI is derived using the pre-trained random forest models, the quality flag will be used to determine the valid data for retrieval, while the rest of the grids will be assigned with fill values. The biome type and land water mask are used as auxiliary data to select the biome model and filter the water surface. The LAI physical valid range is defined as 0 to 7, all the grid with out-of-range data will be marked as the invalid retrieval.

The consistency between training dataset input and operational input data is critical to the algorithm performance, to identify if the input surface reflectance behaves its general biome characteristics, a convex hull check is performed using the predefined polygons for each biome, this check results will be recorded for further reference.

LAI is a vegetation attribute that undergoes gradual and continuous changes over time, a short period of time compositing will significantly improve the data quality. The weekly compositing is to generate 8-day period LAI, which will select the best quality data as the weekly value, in most case the contaminated observation will lead to a lower vegetation metrics (i.e. VI, FPAR, LAI), here, the highest LAI within a certain compositing 8-day window will be selected as the weekly value. Figure 6 illustrates the process flow of the

weekly LAI. The Quality flag of the weekly LAI will be a replication from the corresponding daily LAI.

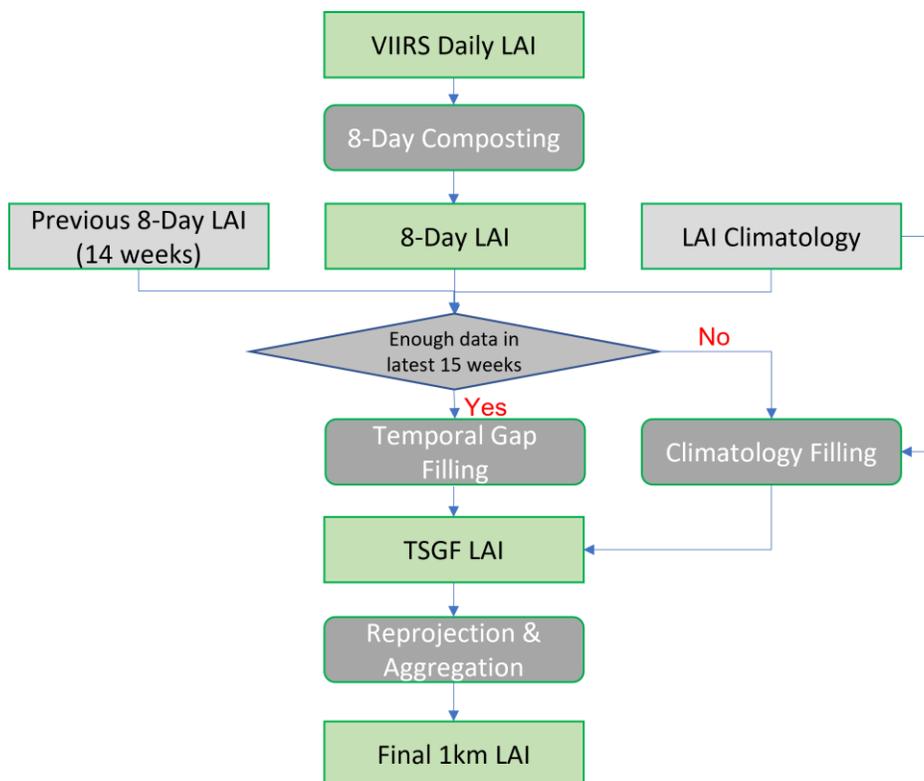


Figure 6 - LAI weekly processes based on the daily retrieval

For this version, the refresh rate will be 8 days, to keep the product date consistent over years, the first compositing window of year will always start from January 1st, and the last compositing window (46th) from day of the year 361 will need the first three or two (for the leap year) days of next year to complete an 8-day window. In the future version, depending on the user request, a daily rolling weekly compositing scheme might be applied to improve the refresh rate.

2.2.5 Temporal smoothing and gap filling.

The weekly compositing will significantly reduce the uncertainty of the daily retrieved LAI, however, the data missing, and outliers are still possible in some situations like cloud, aerosol contamination, darkness or BRDF effect impact. A smoothed vegetation time series is intrinsic feature of the nature, which could be used as an constrain to further improve the satellite derived product, and for the model application, a gap-free product is required, so compositing and temporal smoothing and gap filling (TSGF) will be performed.

For the temporal smoothing process, up to 15 weekly (8-day interval) LAI values will be utilized. Initially, any missing LAI values are replaced by linear interpolation using

adjacent points. Following this, a Savitzky-Golay filter with a window size of 3 is applied to preliminarily smooth the data. Subsequently, an adapted Savitzky-Golay filter with a larger window size of 7 is employed for further smoothing.

Given that LAI is produced as a near real-time product, the temporal series can only incorporate prior data, not post-event data, which poses a significant challenge to the process. For instances lacking sufficient nearby data, climatology data is employed for reconstruction. The available data are utilized to calculate adjustment coefficients (scale factor and offset) in comparison with the climatology curve. Then, the targeted LAI is adjusted based on these climatology values.

In certain scenarios where sufficient data for smoothing are not available (e.g., polar nights during winter, or consistently cloudy conditions in rainforests), climatology data is instead used to ensure data completeness.

Algorithm Input

The input data sets used by the VIIRS LAI algorithm are listed in Tables 5 and 6. The dynamic input data are used in their original swath projection format which is (6400 * 1538 pixels) for the image bands.

Table 5 - List of dynamic input data sets used in the VIIRS LAI algorithm

Input	Datasets
VIIRS Imagery Band Geolocation Data	<ul style="list-style-type: none"> ● Latitude ● Longitude ● Satellite Azimuth Angle ● Satellite Zenith Angle ● Solar Azimuth Angle ● Solar Zenith Angle ● Quality Flag
VIIRS Surface Reflectance EDR	<ul style="list-style-type: none"> ● Channel I1 SR ● Channel I2 SR ● Channel I3 SR ● Quality Flags

The VIIRS surface reflectance is a granule file that contains top of canopy reflectance data in the twelve VIIRS spectral bands listed in Table 2. The LAI algorithm uses the red (I1), NIR (I2) and SWIR (I3) reflectance data. Cloud, aerosol, snow and other quality flag information is included in the files. Details of VIIRS surface reflectance EDR are available at

https://www.star.nesdis.noaa.gov/jpss/documents/ATBD/ATBD_SurfaceReflectance.pdf

The seven bytes of quality flag information included in the VIIRS surface reflectance files are described in the VIIRS Surface Reflectance External User's Manual (EUM) .

The quality of Surface Reflectance (SR) is crucial for the accuracy of LAI retrieval. The VIIRS SR product is directly heritage from collection 5 MODIS and integrated the refinement of the latest Collection 6, currently, the product reach validated maturity stage, indicating that its performance has been thoroughly demonstrated across a broad and diverse range of representative conditions. Validation efforts at AERONET (Aerosol Robotic Network) sites reveal that bands I1, I2, and I3 exhibit high accuracy and precision (as illustrated in Figure 7), fulfilling the product requirements under clear sky conditions. However, there is still potential for improvement in handling challenging atmospheric conditions, which is a primary factor contributing to LAI outliers.

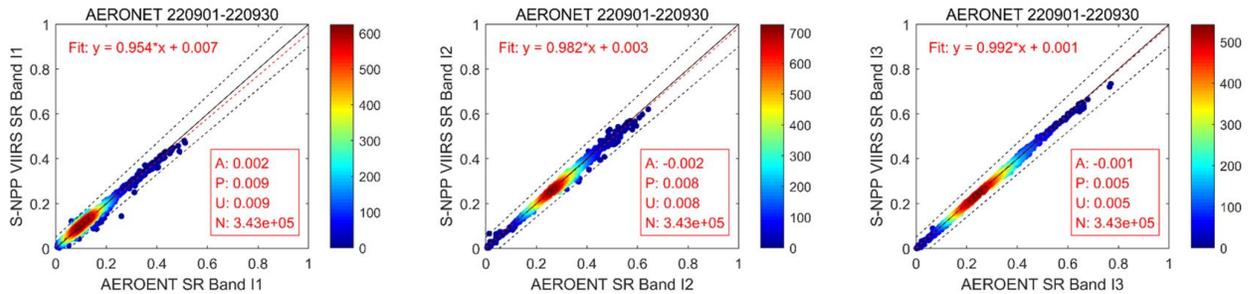


Figure 7 - Surface Reflectance validation results at AERONET

The VIIRS geolocation files (GITCO) include latitude, longitude and solar, view geometry information corresponding to the VIIRS Surface Reflectance. The LAI algorithm uses the latitude and longitude information in gridding and the solar and view geometry information in compositing and retrieval. Further information about the VIIRS geolocation data is available at

https://www.star.nesdis.noaa.gov/ipss/documents/ATBD/D0001-M01-S01-003_JPSS_ATBD_VIIRS-SDR_E.pdf

In the previous section, we covered the preparation of static input data. All these global data are provided in NetCDF format for each tile at 500m resolution, biome type data and land water mask are combined in one file, while LAI climatology is provided as time series data with the same step of 8 days. The biome polygons are archived as six binary matrices measuring 101x101 with a reflectance step increment of 0.01, raster format is more effective for the point in polygon check. The model's coefficients are stored in the KPL format utilizing the Python scikit-learn module, which will also facilitate their loading during the retrieval process. To further streamline the reprojection process, pre-calculated projection indices are supplied, facilitating swift conversion from sinusoidal to equal latitude/longitude.

Table 6 - List of static input used in the VIIRS LAI algorithm

Input	Data Sets
Trained models	Model coefficients for 6 biome types
Biome definition coefficients	Polygon definition matrix
LAI climatology	LAI climatology with same spatial and temporal resolution

Biome type	Biome type data used in model training
Land water mask	Land & water mask
Projection indices	Mapping indices for prompt projection conversion

Algorithm Output

The LAI product is designed with a global scope, featuring a 1km resolution. Originally all the processes are based on a sinusoidal tile grid, the final LAI data undergoes reprojection to fit into a global, equal latitude/longitude grid. This step includes an aggregation process, converting the original 500m resolution to 1km across the globe.

The datasets resulting from this process are detailed in Table 7. The finalized LAI data is offered as a global dataset with a spatial resolution of 0.009 degrees. This data is updated every 8 days, corresponding to a cycle of 46 periods per year, each period spanning 8 days.

Within the data processing workflow, both Daily and Weekly LAI datasets serve as intermediate stages. These datasets are preserved for subsequent analytical processes. Specifically, Daily LAI records are maintained for a duration of 8 days to facilitate the creation of weekly composite data. Meanwhile, Weekly LAI datasets are retained for 15 weeks to enable a temporal smoothing process, enhancing the accuracy and utility of the final product.

Table 7 - VIIRS SR output science datasets

Data Set Name	Description	Dimension
Global LAI	LAI final product	[20000x40000]
Daily LAI	Intermediate data	[2400x2400] for each tile
Weekly LAI	Intermediate data	[2400x2400] for each tile

The LAI valid value is from 0 to 7, adjusted by a scaling factor of 0.01 and an offset value of 0. Quality metrics are generated and retained for oversight purposes and to serve as a reference for users. Unlike direct outputs from the retrieval algorithm, the final LAI product undergoes temporal smoothing. This process involves adjustments based on time series data, rendering the original observation condition or quality flags inapplicable. Instead, a generalized retrieval number and an optimal retrieval percentage are computed for ongoing monitoring purposes. However, for the model application, the product uncertainty is critical to the data assimilation, the grid level uncertainty will be quantified in the future version.

The VIIRS LAI data is available in the NetCDF format, The files adhere to a specific naming convention designed to convey essential information briefly:

WKL-LAI-GLB_v1r0_n20_s202308290000000_e202309052359599_c202403211250073.nc

Product short name: 'WKL-LAI-GLB'

Product version: v1r0

Satellite ID: n20 (For NOAA-20, alternatively 'npp' or 'n21' for SNPP and NOAA-21)

Start time stamp: _sYYYYMMDD0000000 (with fixed starting time 0000000)

End time stamp: _eYYYYMMDD2359599 (with fixed ending time 2359599)

Creation time stamp: _c YYYYMMDDHHMMSSS

Data format: nc (NetCDF)

Performance Estimates

The algorithm theoretical evaluation takes place during the machine learning model's training stage, where it is tested using a portion of the dataset (testing set) that is independent of the training set. This step validates the model's performance from a theoretical standpoint, ensuring that the evaluation accurately reflects the algorithm's capability to generalize to new, unseen data. Specifically, in the case of random forest regression, the result emerges as the average prediction across all trees within the forest. This averaging technique enhances accuracy beyond what any individual tree could achieve, bolstering the algorithm's likelihood of performing well on new data.

In practice, the operational performance of the LAI data may be influenced by several factors beyond algorithmic uncertainty. These include the quality of the input data, which can be affected by variables such as solar and viewing geometry, atmospheric conditions, and the homogeneity of the surface. Additionally, the need to smooth or fill data due to missing values or outliers can further complicate the performance assessment. Post launch validation is for a deep understanding of product performance across all natural conditions, then supporting the refinement of algorithms and uncertainty estimation for users.

2.5.1 in-situ LAI validation

The primary in-situ validation ground datasets are provided by the Ground-Based Observations for Validation (GBOV) service, which were conducted at 19 distinct sites within the National Ecology Observatory Network (NEON) across the continental United States. The ground LAI were acquired using digital hemispherical photography (DHP), a technique involving the capture of both upward-facing and downward-facing photos in accordance with the NEON sampling protocol. This method effectively captures the canopy structure, aiding in accurate LAI assessment. The raw DHP datasets are then processed with great care before being made accessible via the GBOV service.

It is important to note that the indirect measurement method primarily derives the Plant Area Index (PAI) as a surrogate for LAI during validation processes. PAI includes both leafy and woody elements within its measurements. Therefore, in forested areas where woody components are more prevalent, PAI estimates tend to be higher. To address this, a correction process is implemented to subtract the woody part from PAI, yielding a more

accurate representation of LAI. Additionally, adjustments are made to account for the clumping effect, which involves converting the effective PAI to a true PAI.

Moreover, the in-situ reference measurements were upscaled using high-resolution (10-30m) satellite data from sources such as Sentinel-2 and Landsat. This upscaling was essential to align these ground measurements with LAI products derived from satellites with moderate spatial resolution. Figure 8 demonstrates the upscaled LAI reference data, with the 30m Landsat reflectance data, high resolution LAI is retrieved using an ANN algorithm, then multiple temporal ground LAI are used for the LAI calibration then get the 30m LAI reference map, for VIIRS validation, aggregation process is necessary to match the spatial resolution.

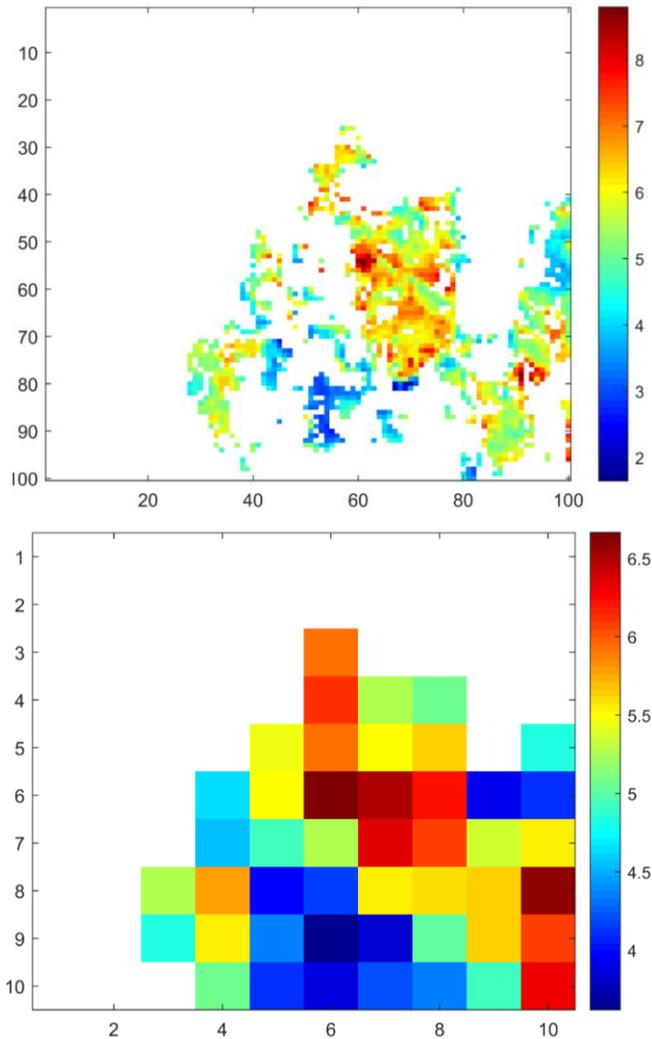


Figure 8 - The ground LAI upscaling using the high-resolution LAI (Left: 30m Landsat LAI, Right: aggregated 300m LAI) as bridge

More details about the ground measurement methodologies, the study by Meier et al. (2018) offers comprehensive insights into NEON's protocols. The processing techniques

and distribution strategies utilized by GBOV are extensively detailed in Brown et al. (2020). Further information is readily available on the GBOV website (<https://gbov.acri.fr>).

For this version, the refresh rate will be 8 days, to keep the product date consistent over years, the first compositing window of year will always start from January 1st, and the last compositing window (46th) from day of the year 361 will need the first three or two (for the leap year) days of next year to complete an 8-day window. In the future version, depending on the user request, a daily rolling weekly compositing scheme might be applied to improve the refresh rate.

Preliminary validation of the LAI was performed using one year of ground data supplied by GBOV. Figure 9 showcases scatter plots comparing the ground-based LAI measurements with the VIIRS LAI product data. Additionally, the figure summarizes the validation statistics. A good correlation between the ground LAI and VIIRS LAI product was observed, demonstrating acceptable levels of accuracy and precision. However, there are still some grids with uncertainties exceeding the required thresholds. This suggests that both the LAI product and the validation process require further refinement.

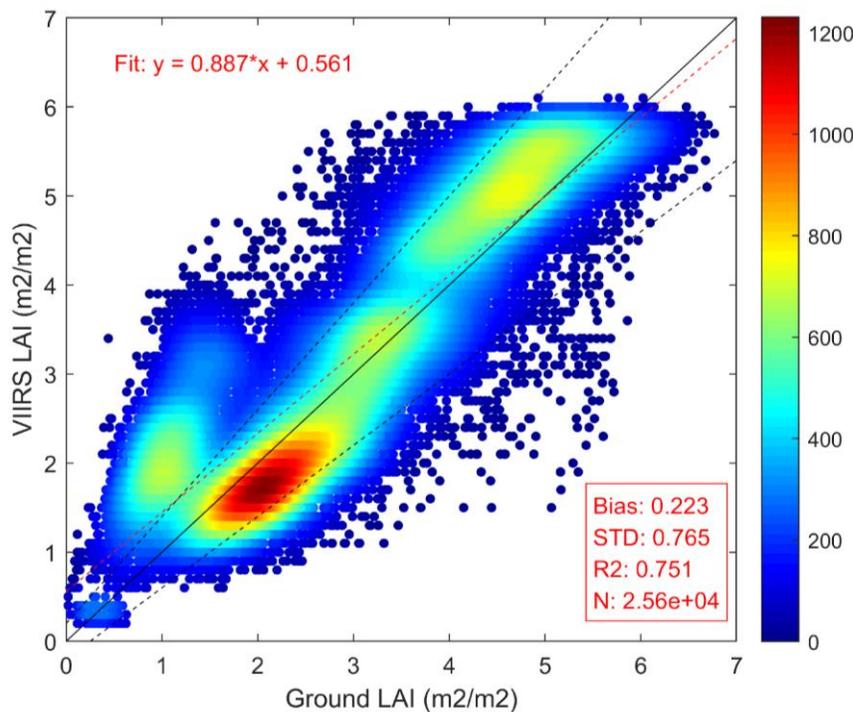


Figure 9 - The LAI product preliminary validation at NEON sites

To enhance the robustness of the LAI product validation, data from other field campaigns will also be employed, supplementing the measurements from the NEON sites. This approach is taken to ensure the validation process is comprehensive and not biased by

a limited dataset. However, it's important to acknowledge that some of the data from these campaigns were collected in previous years. Consequently, there is a need to reprocess the VIIRS Surface Reflectance data to generate LAI estimates that are consistent and suitable for validation purposes.

2.5.2 Inter-comparison with existing products

To evaluate the consistency between the newly developed LAI product and established LAI training products, we utilized sites from the Benchmark Land Multisite Analysis and Intercomparison of Products (BELMANIP) network, version 2.1, for site-scale comparisons. BELMANIP2, an evolution of the original network, integrates sites from renowned experimental networks—including FLUXNET, AERONET, VALERI, and BigFoot—augmented by strategically chosen locations from the GLC2000 land cover map. The expanded BELMANIP2 dataset now includes 445 diverse sites, carefully selected to reflect the vast array of global vegetation types and climatic conditions. This deliberate selection process ensures that, despite the dataset's size, it provides a solid and representative basis for conducting meaningful comparative analyses between the LAI products. Leveraging the carefully curated BELMANIP dataset allows for an effective and comprehensive assessment, ensuring a wide-ranging representation in the comparison of LAI products.

As the primary dataset for training, the NASA VNP15 LAI product is utilized for inter-comparison, with the expectation that the newly developed LAI will perform comparably to VNP15. This assumption is validated through cross-comparison, which, supported by high-quality data spanning an entire year, demonstrates good agreement between the two products, as illustrated in Figure 10. The figure also includes relevant statistical analysis. However, the new LAI product offers certain advantages over VNP15. During the preparation of the training data, two widely used products, GLASS and GEOv2, are integrated to filter out questionable data. The GLASS LAI, benefiting from time series information for offline processing, provides smooth data. Meanwhile, GEOv2, being an independent observation distinct from MODIS/VIIRS data, contributes to the fused LAI used for model training, enhancing its reliability. Additionally, the temporal smoothing process significantly reduces outliers. Although this process may introduce uncertainty in certain scenarios, it enhances data usability and overall quality for a global gap-free product.

A similar inter-comparison will also be conducted for the GLASS and GEOv2 products alongside the new product. This process aims to identify the strengths of these existing products and incorporate them into the new LAI product. Additionally, it seeks to address any shortcomings they may have. This comparative analysis will facilitate a comprehensive understanding of how the new LAI product can leverage the advantages of the widely used products while striving to mitigate their limitations, ultimately leading to the development of a more refined and effective LAI product.

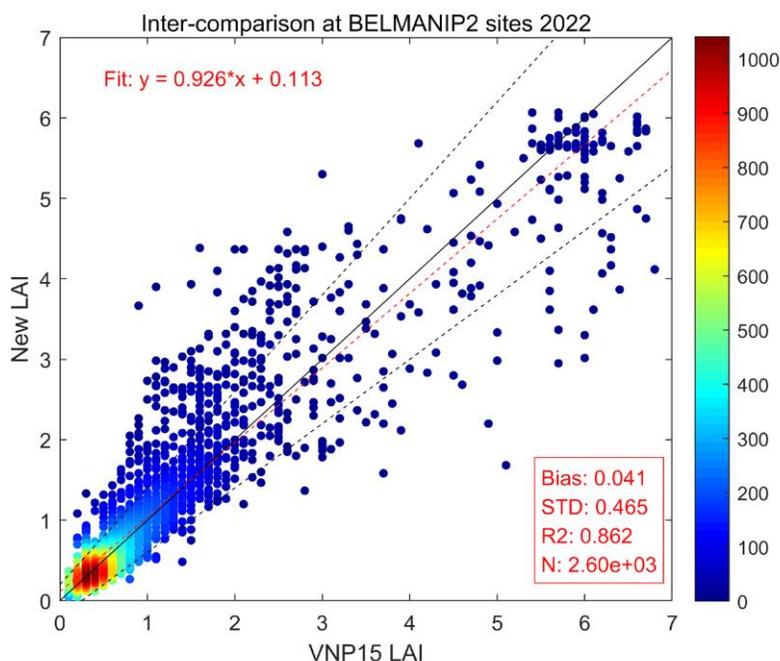


Figure 10 - Preliminary inter-comparison with VNP15 LAI at the BELMANIP2 sites

The BELMANIP dataset can provide a general overview of the global performance of these LAI products. However, to conduct a more detailed analysis, it is necessary to examine the LAI distribution across different biome types specifically.

Beyond absolute values, the temporal continuity and smoothness of the LAI are crucial metrics for evaluating the product's quality. To assess the temporal smoothness, a simple Temporal Discontinuity Index (TDI) will be employed. The TDI is defined as the mean of the absolute differences between all adjacent time steps in the LAI series, represented by the formula:

$$TDI = \sum_{t=1}^{N_t-1} |LAI_{t+1} - LAI_t|$$

In this formula, N_t denotes the number of time steps for a given period, while LAI_t and LAI_{t+1} are the LAI values at adjacent time points in the series. This index will provide a quantitative measure of how smoothly the LAI values transition over time, highlighting the effectiveness of temporal smoothing processes and the overall continuity of the product.

2.5.3 LAI performance in Noah-MP model

The Noah-MP LSM is an advanced land surface modeling framework that represents physical and biophysical exchanges between the atmosphere and land surface, including energy, water, and carbon fluxes. It is slated to be the next-generation LSM for numerical

weather forecasting. The Noah-MP LSM enhances the original Noah LSM by offering multiple parameterization options for key processes, with LAI playing a pivotal role in these simulations.

Given Noah-MP will be the primary user, the LAI impact on the model is an import metric to evaluate its performance. The current LSM model taking prescribed LAI based on a monthly schedule for each of 28 biome types, the Noah-MP test will integrate LAI data progressively, Initially, global LAI monthly climatology will be incorporated into the model, followed by an assessment of its impact on model outputs. Should the inclusion of LAI data demonstrate a positive effect, a dynamic LAI dataset will then be introduced for further evaluation.

To enhance model support, a deep understanding of vegetation physical mechanisms and parameterization schemes is essential for applying LAI, especially over heterogeneous surfaces where a single LAI value may not sufficiently capture vegetation dynamics. Efforts should be made to explore effective methods for integrating current LAI products into the model, thereby improving its accuracy and reliability.

ASSUMPTIONS AND LIMITATIONS

Performance Assumptions

The LAI retrieval algorithm performance is dependent on observation conditions and the input data consistency and quality.

- VIIRS surface reflectance in I1, I2 and I3 bands are available, calibrated and navigated and are quality assured.
- The training datasets are consistent with the operational input data (surface reflectance, biome type)

The final LAI is temporally smoothed, and gap filled, the following assumptions should be satisfied:

- The assumption that LAI changes slowly and continuously from week to week, and within a week (8 days period) no dramatic LAI change.
- The biome type is stable over time.

Potential Improvements

For machine learning algorithms, product performance is closely tied to the quality and representativeness of the training data. As such, updates to the machine learning model are made with enhancements to the training dataset, driven by the evaluation and validation results from operational data. This iterative process ensures continuous improvement and relevance of the algorithm to its applications.

In the context of near real-time LAI temporal smoothing, a significant challenge arises due to the availability of only prior information in the time series. This limitation becomes

particularly problematic during periods of rapid vegetation changes, where the absence of post-event data can introduce substantial uncertainty. While leveraging LAI climatology can mitigate this to some extent, it is recognized that future enhancements will likely involve more sophisticated time series analysis methods to address these challenges effectively.

The LAI product is specifically tailored to support the Noah-MP LSM, recognizing the importance of this integration, potential future improvements or adjustments are anticipated to ensure the product's efficient compatibility and performance within the model framework. This forward-looking approach underscores the commitment to refining the LAI product in response to evolving model requirements and scientific advancements.

REFERENCES

- Arsenault, K.R., Nearing, G.S., Wang, S., Yatheendradas, S. and Peters-Lidard, C.D., 2018. Parameter sensitivity of the Noah-MP land surface model with dynamic vegetation. *Journal of Hydrometeorology*, 19(5), pp.815-830.
- Baret, F., Morisette, J.T., Fernandes, R.A., Champeaux, J.L., Myneni, R.B., Chen, J., Plummer, S., Weiss, M., Bacour, C., Garrigues, S. and Nickeson, J.E., 2006. Evaluation of the representativeness of networks of sites for the global validation and intercomparison of land biophysical products: Proposition of the CEOS-BELMANIP. *IEEE Transactions on Geoscience and Remote Sensing*, 44(7), pp.1794-1803.
- Baret, F., Hagolle, O., Geiger, B., Bicheron, P., Miras, B., Huc, M., et al. (2007). LAI, fPAR, and fCover CYCLOPES global products derived from VEGETATION part 1: Principles of the algorithm. *Remote Sensing of Environment*, 110(3), 275–286.
- Baret, F., Weiss, M., Lacaze, R., Camacho, F., Makhmara, H., Pacholczyk, P., & Smets, B. (2013). GEOV1: LAI, FAPAR essential climate variables and FCOVER global time series capitalizing over existing products. Part1: Principles of development and production. *Remote Sensing of Environment*, 137, 399–309.
- Brown, L., Chen, J.M., Leblanc, S.G. and Cihlar, J., 2000. A shortwave infrared modification to the simple ratio for LAI retrieval in boreal forests: An image and model analysis. *Remote sensing of environment*, 71(1), pp.16-25.
- Brown, L.A., Meier, C., Morris, H., Pastor-Guzman, J., Bai, G., Lerebourg, C., Gobron, N., Lanconelli, C., Clerici, M. and Dash, J., 2020. Evaluation of global leaf area index and fraction of absorbed photosynthetically active radiation products over North America using Copernicus Ground Based Observations for Validation data. *Remote Sensing of Environment*, 247, p.11935.
- Camacho, F., Cernicharo, J., Lacaze, R., Baret, F. and Weiss, M., 2013. GEOV1: LAI, FAPAR essential climate variables and FCOVER global time series capitalizing over existing products. Part 2: Validation and intercomparison with reference products. *Remote Sensing of Environment*, 137, pp.310-329.
- Camacho, F., Lacaze, R., Latorre, C., Baret, F., De la Cruz, F., Demarez, V., Di Bella, C., García-Haro, J., González-Dugo, M.P., Kussul, N. and Mattar, C., 2015. Collection of ground biophysical measurements in support of copernicus global land product validation: The imagines database. *measurements*, 17, p.2209.
- Cao, C., F. DeLuccia, X. Xiong, R. Wolfe, F. Weng, Early On-orbit Performance of the Visible Infrared Imaging Radiometer Suite (VIIRS) onboard the Suomi National Polar-

orbiting Partnership (S-NPP) Satellite, IEEE Trans. on Geosci. and Remote Sens., 2014, Page(s): 1142 - 1156.

Cao, C., Zhang, B., Shao, X., Wang, W., Uprety, S., Choi, T., Blonski, S., Gu, Y., Bai, Y., Lin, L. and Kalluri, S., 2021. Mission-long recalibrated science quality Suomi NPP VIIRS radiometric dataset using advanced algorithms for time series studies. Remote Sensing, 13(6), p.1075.

Chen, J., Jönsson, P., Tamura, M., Gu, Z., Matsushita, B. and Eklundh, L., 2004. A simple method for reconstructing a high-quality NDVI time-series data set based on the Savitzky–Golay filter. Remote sensing of Environment, 91(3-4), pp.332-344.

Chen, J.M., Menges, C.H. and Leblanc, S.G., 2005. Global mapping of foliage clumping index using multi-angular satellite data. Remote Sensing of Environment, 97(4), pp.447-457.

Claverie, M., Matthews, J.L., Vermote, E.F. and Justice, C.O., 2016. A 30+ year AVHRR LAI and FAPAR climate data record: Algorithm description and validation. Remote Sensing, 8(3), p.263.

Fang, H., Jiang, C., Li, W., Wei, S., Baret, F., Chen, J.M., Garcia-Haro, J., Liang, S., Liu, R., Myneni, R.B. and Pinty, B., 2013. Characterization and intercomparison of global moderate resolution leaf area index (LAI) products: Analysis of climatologies and theoretical uncertainties. Journal of Geophysical Research: Biogeosciences, 118(2), pp.529-548.

Fang, H., Baret, F., Plummer, S. and Schaepman-Strub, G., 2019. An overview of global leaf area index (LAI): Methods, products, validation, and applications. Reviews of Geophysics, 57(3), pp.739-799.

Fernandes, R., Plummer, S., Nightingale, J., Baret, F., Camacho, F., Fang, H., Garrigues, S., Gobron, N., Lang, M., Lacaze, R., LeBlanc, S., Meroni, M., Martinez, B., Nilson, T., Pinty, B., Pisek, J., Sonnentag, O., Verger, A., Welles, J., Weiss, M., & Widlowski, J.L. (2014). Global Leaf Area Index Product Validation Good Practices. Version 2.0.

Houborg, R. and McCabe, M.F., 2018. A hybrid training approach for leaf area index estimation via Cubist and random forests machine-learning. ISPRS Journal of Photogrammetry and Remote Sensing, 135, pp.173-188.

Jacquemoud, S., Verhoef, W., Baret, F., Bacour, C., Zarco-Tejada, P.J., Asner, G.P., François, C. and Ustin, S.L., 2009. PROSPECT+ SAIL models: A review of use for vegetation characterization. Remote sensing of environment, 113, pp.S56-S66.

Jiang, Z., Vargas, M. and Csaszar, I., 2016, July. New operational real-time daily rolling weekly Green Vegetation fraction product derived from suomi NPP VIIRS reflectance data. In 2016 IEEE International Geoscience and Remote Sensing Symposium (IGARSS) (pp. 3524-3527). IEEE.

Kang, Y., Özdoğan, M., Zipper, S.C., Román, M.O., Walker, J., Hong, S.Y., Marshall, M., Magliulo, V., Moreno, J., Alonso, L. and Miyata, A., 2016. How universal is the relationship between remotely sensed vegetation indices and crop leaf area index? A global assessment. *Remote sensing*, 8(7), p.597.

Kang, Y., Ozdogan, M., Gao, F., Anderson, M.C., White, W.A., Yang, Y., Yang, Y. and Erickson, T.A., 2021. A data-driven approach to estimate leaf area index for Landsat images over the contiguous US. *Remote Sensing of Environment*, 258, p.112383.

Kumar, Sujay V., David M. Mocko, Shugong Wang, Christa D. Peters-Lidard, and Jordan Borak. "Assimilation of Remotely Sensed Leaf Area Index into the Noah-MP Land Surface Model: Impacts on Water and Carbon Fluxes and States over the Continental United States." *Journal of Hydrometeorology* 20, no. 7 (2019): 1359-1377.

Liu, X., Chen, F., Barlage, M., Zhou, G. and Niyogi, D., 2016. Noah-MP-Crop: Introducing dynamic crop growth in the Noah-MP land surface model. *Journal of Geophysical Research: Atmospheres*, 121(23), pp.13-953.

Ma, H. and Liang, S., 2022. Development of the GLASS 250-m leaf area index product (version 6) from MODIS data using the bidirectional LSTM deep learning model. *Remote sensing of environment*, 273, p.112985.

Meier, C., Everhart, J. and Jones, K., 2018. TOS Protocol and Procedure: Measurement of Leaf Area Index, K. ed. National Ecological Observatory Network, Boulder, Colorado, United States.

Miura, T., Muratsuchi, J. and Vargas, M., 2018. Assessment of cross-sensor vegetation index compatibility between VIIRS and MODIS using near-coincident observations. *Journal of Applied Remote Sensing*, 12(4), pp.045004-045004.

Myneni, R. B., Hoffman, S., Knyazikhin, Y., Privette, J. L., Glassy, J., Tian, Y., et al. (2002). Global products of vegetation leaf area and fraction absorbed PAR from year one of MODIS data. *Remote Sensing of Environment*, 83(1-2), 214–231.

Niu, G.Y., Yang, Z.L., Mitchell, K.E., Chen, F., Ek, M.B., Barlage, M., Kumar, A., Manning, K., Niyogi, D., Rosero, E. and Tewari, M., 2011. The community Noah land surface model with multiparameterization options (Noah-MP): 1. Model description and evaluation with local-scale measurements. *Journal of Geophysical Research: Atmospheres*, 116(D12).

Roger, J., Vermote, E., and Ray, J., 2015. MODIS Surface Reflectance User's Guide, Collection 6.

Savitzky, A. and Golay, M.J., 1964. Smoothing and differentiation of data by simplified least squares procedures. *Analytical chemistry*, 36(8), pp.1627-1639.

Secretariat, G.C.O.S., 2006. Systematic observation requirements for satellite-based products for climate. GCOS Implementation Plan.

Vermote, E., Justice, C. and Csiszar, I., 2014. Early evaluation of the VIIRS calibration, cloud mask and surface reflectance Earth data records. *Remote Sensing of Environment*, 148, pp.134-145.

Verger, A., Baret, F., Weiss, M., Kandasamy, S. and Vermote, E., 2013. The CACAO method for smoothing, gap filling, and characterizing seasonal anomalies in satellite time series. *IEEE transactions on Geoscience and Remote Sensing*, 51(4), pp.1963-1972.

Verger, A., Baret, F. and Weiss, M., 2014. Near real-time vegetation monitoring at global scale. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 7(8), pp.3473-3481.

Verger, A., Sánchez-Zapero, J., Weiss, M., Descals, A., Camacho, F., Lacaze, R. and Baret, F., 2023. GEOV2: Improved smoothed and gap filled time series of LAI, FAPAR and FCover 1 km Copernicus Global Land products. *International Journal of Applied Earth Observation and Geoinformation*, 123, p.103479.

Xiao, Z., Liang, S., Wang, J., Chen, P., Yin, X., Zhang, L., & Song, J. (2014). Use of general regression neural networks for generating theGLASS leaf area index product from time-series MODIS surface reflectance. *IEEE Transactions on Geoscience and Remote Sensing*, 52(1),209–223.

Xu, B., Li, J., Park, T., Liu, Q., Zeng, Y., Yin, G., Zhao, J., Fan, W., Yang, L., Knyazikhin, Y. and Myneni, R.B., 2018. An integrated method for validating long-term leaf area index products using global networks of site-based measurements. *Remote Sensing of Environment*, 209, pp.134-151.

Yan, K., T. Park, G. Yan, Z. Liu, B. Yang, C. Chen, R. Nemani, Y. Knyazikhin, and R.B. Myneni. 2016b. "Evaluation of MODIS LAI/FPAR Product Collection 6. Part 2: Validation and Intercomparison." *Remote Sensing* 8: 6. doi: 10.3390/rs8060460.

Yan, K., Park, T., Chen, C., Xu, B., Song, W., Yang, B., et al. (2018). Generating global products of LAI and FPAR from SNPP-VIIRS Data:Theoretical background and

implementation. IEEE Transactions on Geoscience and Remote Sensing, 56(4), 2119–2137.

Yin, G., Li, A., Jin, H., Zhao, W., Bian, J., Qu, Y., Zeng, Y. and Xu, B., 2017. Derivation of temporally continuous LAI reference maps through combining the LAINet observation system with CACAO. Agricultural and Forest Meteorology, 233, pp.209-221.

Yuan, H., Dai, Y., Xiao, Z., Ji, D. and Shanguan, W., 2011. Reprocessing the MODIS Leaf Area Index products for land surface and climate modelling. Remote Sensing of Environment, 115(5), pp.1171-1187.

Zhang, R., Huang, C., Zhan, X., Jin, H. and Song, X.P., 2018. Development of S-NPP VIIRS global surface type classification map using support vector machines. International Journal of Digital Earth, 11(2), pp.212-232.

Zheng, W., Wei, H., Wang, Z., Zeng, X., Meng, J., Ek, M., Mitchell, K. and Derber, J., 2012. Improvement of daytime land surface skin temperature over arid regions in the NCEP GFS model and its impact on satellite data assimilation. Journal of Geophysical Research: Atmospheres, 117(D6).

END OF DOCUMENT