NOAA Leaf Area Index Climate Data Record

1. Intent of This Document and Point of Contact

1a) This document is intended for users who wish to compare observation-based data-products with climate model output in the context of the CMIP/IPCC historical experiments. Users are not expected to be experts in observation-based data. This document summarizes essential information needed for comparing this dataset to climate model output. References are provided at the end of this document to additional information.
Dataset File Names (as appears on the ESGF):
obs4MIPs.NOAA-NCEI.NOAA-NCEI-LAI-AVHRR-5-0.mon.lai.gn
obs4MIPs.NOAA-NCEI.NOAA-NCEI-LAI-VIIRS-1-0.mon.lai.gn

1b) Technical point of contact for this dataset who is able to field questions from users about all aspects of the dataset:
Eric Vermote, NASA Goddard Space Flight Center (GSFC), eric.f.vermote@nasa.gov
Jessica Matthews, NOAA's National Centers for Environmental Information (NCEI), jessica.matthews@noaa.gov

2. Data Field Description

CF standard variable name, if applicable, units:	*leaf_area_index (CF standard name), 1* *lai (CMIP variable), 1*
Long name for variable	*Leaf Area Index (LAI) is a dimensionless quantity that represents the total one-sided leaf area per unit of ground surface area.*
Spatial resolution:	*0.05-deg in latitude and longitude*
Temporal resolution and extent:	*Monthly average, from 06/1981 to 12/2024*
Coverage:	*Global*

3. Data Origin

This dataset uses the Surface Reflectance - Polar Orbiter climate data record (CDR) as the primary input. Primary sensor data for the surface reflectance data are calibrated and geolocated AVHRR Global Area Coverage (GAC) 'Level 1B' reflectance, the VIIRS 375-m Earth View SDR, geolocation information, and sensor data quality flags. Surface Reflectance products are generated for each cloud-free pixel (0.05°x0.05°) observed by the AVHRR imager channels 1 and 2 and VIIRS channels I1 and I2. Other channels are used to retrieve atmospheric conditions, (e.g., water vapor) cloud and snow masks.

Multiple ancillary data sources are used in combination with the primary AVHRR GAC and VIIRS SDR inputs to produce the surface reflectance data. Among the inputs are digital elevation models from USGS, land/water masks from NASA's Moderate Resolution Imaging Spectroradiometer (MODIS) and VIIRS, water vapor information from National Centers for Environmental Prediction (NCEP), ozone data from the Ozone Mapping and Profiler Suite (OMPS) or NCEP, the bidirectional reflectance distribution function (BRDF) database from MODIS, BRDF-corrected reflectance climatologies from MODIS, as well as internally derived stratospheric and tropospheric aerosol climatologies.

Various corrections are made to the primary inputs in the derivation of surface reflectance. This includes performing BRDF-correction and atmospheric correction accounting for Rayleigh scattering, stratospheric and tropospheric aerosols, and gaseous absorption of O_3, O_2, CO_2, and water vapor.

Furthermore, pixels with cloud or cloud shadow present are not included in the dataset. This is assessed with comparisons of the spectrally-adjusted AVHRR to the MODIS Channel 1 BRDF-corrected climatology. If the difference between the AVHRR reflectance and the climatology is larger than 0.03, the pixel is considered to be cloud. Cloud height range (minimum and maximum) is estimated based on temperature derived from AVHRR Channel 4 and 5. Shadow pixels correspond to the projection of the cloudy pixels on the surface following the pixel location (lat, lon) – sun angle. For the VIIRS inputs, the VIIRS cloud mask is used.

To compute Leaf Area Index (LAI) from the surface reflectance data, an internal BRDF database and a land cover classification based on Hansen et al. (1998), both with 0.05° spatial resolution, were used as ancillary data. Globally land cover was binned into 6 classes (as shown in Figure 1): water, needleleaf forest, evergreen broadleaf forest, broadleaf forest, shrublands, and grasslands/croplands/non-vegetated.

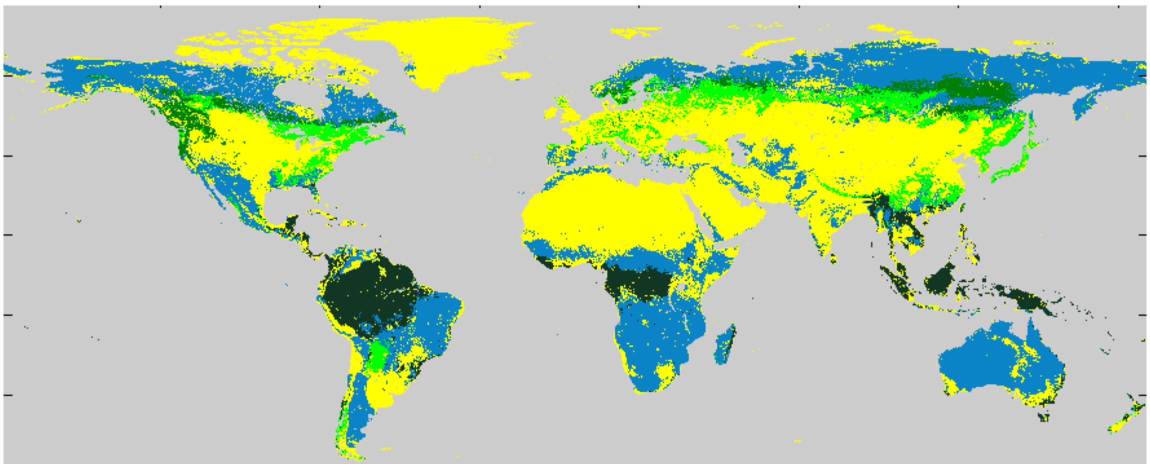

Figure 1: *Landcover for year 2004. Grey: water; blue: shrubland; yellow: Grasslands & Croplands & Non-vegetated; light green: broadleaf forest; green: needleleaf forest; dark green: evergreen broadleaf forest.*

An artificial neural network (ANN) connecting LAI and surface reflectance for each of the 5 land cover biomes was trained using MODIS LAI (aggregated from 1 km to 0.05°, and from 8-day to monthly) and AVHRR surface reflectance data (at 0.05° and monthly) from 2000-2007. The ANN were trained over a defined area and the output accuracy decreases considerably outside of the domain delimited by the learning dataset. Therefore, an acceptable input domain is defined for each class based on surface reflectance inputs used during ANN training. Figure 2 illustrates the density distribution of the learning dataset for each class and the associated domain is delimited by a polygon. Polygons were defined to include 97% of the density distribution pixels (0.01 resolution for red and NIR).

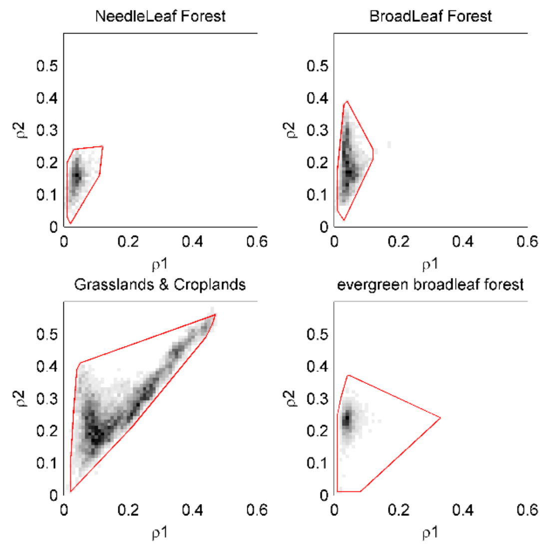

Figure 2: *Domain definition for the 5 classes (red polygons) in the red/NIR surface reflectance space. Greyscale images represent the density function for each 0.01 surface reflectance bin (white = no value; black = high density).*

4. Validation and Uncertainty Estimate

Direct comparison of the satellite-derived products with *in situ* measurements is a key validation step. However, an important issue related to the validation of any coarse resolution retrieval is to link the pixel footprint to the spatial representativeness of the measurement. We relied on the work performed by Garrigues et al. (2008) who contributed to the conception of the DIRECT network. DIRECT is a collection of sites for which ground measurements have been collected and processed according to the CEOS-LPV (Centre for Earth Observation Science—Land Product Validation) guidelines. They first gathered in situ measurements from many locations and scaled them up to a 3 × 3 km area using medium-resolution (<100 m) data. To extend the measurement from a 3 km × 3 km area to a 0.05° area, we applied a ratio calculated using the 1 km MODIS LAI retrieval aggregated over the measurement footprint (3 km × 3 km) and the one aggregated at 0.05°. The outputs were finally compared to the LAI CDR retrieval (Figure 3). LAI validation scatter plots were separated among the type of measurement: effective and true LAI, depending if the clumping factor is considered or not (Claverie, 2012; Claverie et al., 2013).

Three statistical metrics (Equations (1)–(3)) are calculated: accuracy (A), precision (P), and uncertainty (U).

$$A = \frac{1}{N} \sum_{i=1}^{N} \varepsilon_i \tag{1}$$

$$P = \sqrt{\frac{1}{N-1} \sum_{i=1}^{N} (\varepsilon_i - A)^2} \tag{2}$$

$$U = \sqrt{\frac{1}{N} \sum_{i=1}^{N} \varepsilon_i^2} \tag{3}$$

In Equations (1)–(3), *N* is the number of valid samples used for the comparison and ε_i is the estimate minus the reference. Relative values for the three metrics are computed by dividing the metric by the mean value of the reference observation.

The error budget is detailed in Table 1, which includes per-class accuracy, precision and uncertainty from the validation over DIRECT sites. The class “Grasslands and Croplands and Non-vegetated” is the most represented class of the *in situ* dataset, in terms of the largest N. The computed uncertainty, also referred to as RMSD, fits within the range of previously published comparisons. Camacho et al. (2013) validated four global LAI products (GEOV1, CYCLOPE, MCD15, and GLOV2) and found RMSD ranges of 0.74–1.39 for effective LAI.

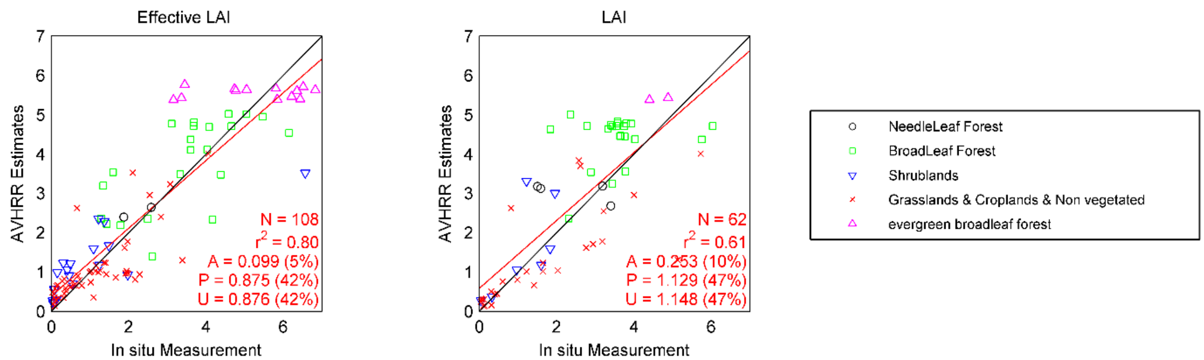

Figure 3: *In situ validation of LAI over DIRECT sites. Ground measurement covers initially a footprint of 3 km × 3 km and were extrapolated to 0.05° using MODIS products for direct comparison. Statistical metrics are defined in Equations (1)–(3); values in parenthesis correspond to metric values divided by the reference mean value.*

Class	Effective LAI				True LAI			
	A	P	U	N	A	P	U	N
NeedleLeaf Forest	0.29	0.32	0.37	2	0.62	1.18	1.19	4
BroadLeaf Forest	0.28	1.03	1.05	22	0.81	1.02	1.29	22
Shrublands	0.2	0.9	0.9	20	0.4	0.88	0.91	7
Grasslands & Croplands & Non-vegetated	-0.08	0.65	0.65	51	-0.33	1.07	1.1	27
Evergreen broadleaf forest	0.14	1.35	1.31	14	0.76	0.3	0.79	2
All	**0.08**	**0.89**	**0.89**	**109**	**0.25**	**1.13**	**1.15**	**62**

Table 1: *Error budget based on in situ validation. N corresponds to the number of points used to compute the statistical metrics.*

5. Considerations for Model-Observation Comparisons

The original LAI CDR contains a wealth of associated quality assurance information (see Table 2). For ease of use in the obs4MIPS context, we chose to only include CDR data under certain conditions for the monthly averages. In particular, we only include data with the "OK" quality flag set. That is, the data is not included if any of the following are indicated: "Input flag is cloudy", "Invalid input", or "Output out of range". Further, the original LAI CDR has a daily temporal resolution. However, the obs4MIPS version has monthly temporal resolution derived by taking the average of all pixels that passed quality filtering for the month.

A known limitation of the algorithm is the capacity to characterize the LAI dynamics for the Evergreen broadleaf forest class. This is due to the saturation of Channels 1 and 2 signals over dense vegetation cover. Another area for future improvement is that the current ancillary land cover classification is produced using only the 1981-1994 time period, land cover should ideally be a dynamic input to account for large scale land cover changes over time (e.g. deforestation).

Bit #	Description	Definition
6-7	Polygon test	00: in polygon
		01: not in polygon
		10: not tested (water/cloudy)
5	BRDF corrected	0: no
		1: yes
2-4	Associated Class	001: NeedleLeaf Forest
		010: BroadLeaf Forest
		011: Shrublands
		100: Grasslands & Croplands & Non-vegetated
		101: Evergreen broadleaf forest
		110: Water
0-1	Quality control	00: OK
		01: Input flag as Cloudy
		10: Invalid input
		11: Output out of range

Table 2: *Quality assurance description. Bits are listed from the most significant bit (bit 7) to the least significant bit (bit 0).*

6. Any Other Relevant Information

N/A

7. Dataset citation and Acknowledgments

https://doi.org/10.25981/ESGF.obs4MIPs.CMIP7/2997962

8. References

8a) Core reference(s), documents and related material for this dataset:

Claverie, M., J.L. Matthews, E.F. Vermote, C.O. Justice. (2016). A 30+ Year AVHRR LAI and FAPAR Climate Data Record: Algorithm description and validation. *Remote Sensing*, 8(3), 263, https://doi:10.3390/rs8030263.

Claverie, Martin; Vermote, Eric; Justice, Chris; Csiszar, Ivan; Myneni, Ranga; Baret, Frederic; Masuoka, Ed; Wolfe, Robert; Ray, James P.; NOAA CDR Program. (2023): NOAA Climate Data Record (CDR) of VIIRS Leaf Area Index (LAI) and Fraction of Absorbed Photosynthetically Active Radiation (FAPAR), Version 1. NOAA National Centers for Environmental Information. https://doi.org/10.25921/9x3m-0e02.

Ray, J. (2023). CDR Program Climate Algorithm Theoretical Basis Document: VIIRS Leaf Area Index (LAI) and Fraction of Absorbed Photosynthetically Active Radiation (FAPAR).

https://www.ncei.noaa.gov/pub/data/sds/cdr/CDRs/Leaf_Area_Index_and_FAPAR/VIIRS/AlgorithmDescriptionVIIRS_01B-20c.pdf

Vermote, Eric; NOAA CDR Program. (2019): NOAA Climate Data Record (CDR) of AVHRR Leaf Area Index (LAI) and Fraction of Absorbed Photosynthetically Active Radiation (FAPAR), Version 5. NOAA National Centers for Environmental Information. https://doi.org/10.7289/V5TT4P69.

Vermote, E. (2018). CDR Program Climate Algorithm Theoretical Basis Document: Leaf Area Index (LAI) and Fraction of Absorbed Photosynthetically Active Radiation (FAPAR). https://www.ncei.noaa.gov/pub/data/sds/cdr/CDRs/Leaf_Area_Index_and_FAPAR/AVHRR/AlgorithmDescriptionAVHRR_01B-20c.pdf

8b) Additional information reference(s), documents and related material for this dataset:

Camacho, F., Cernicharo, J., Lacaze, R., Baret, F., & 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*, 310–329.

Claverie, M. (2012). *Estimation spatialisée de la biomasse et des besoins en eau des cultures à l'aide de données satellitales à hautes résolutions spatiale et temporelle: Application aux agrosystèmes du sud-ouest de la france* [Unpublished doctoral dissertation]. Université Toulouse III Paul Sabatier.

Claverie, M., Vermote, E. F., Weiss, M., Baret, F., Hagolle, O., & Demarez, V. (2013). Validation of coarse spatial resolution LAI and FAPAR time series over cropland in southwest france. **Remote Sensing of Environment**, *139*, 216–230.

Garrigues, S., Lacaze, R., Baret, F., Morisette, J. T., Weiss, M., Nickeson, J. E., Fernandes, R., Plummer, S., Shabanov, N. V., & Myneni, R. B. (2008). Validation and intercomparison of global leaf area index products derived from remote sensing data. **Journal of Geophysical Research: Biogeosciences**, *113*(G2).

Hansen, M., DeFries, R., Townshend, J. R. G., & Sohlberg, R. (1998). *UMD Global Land Cover Classification, 1 Kilometer, 1.0*. Department of Geography, University of Maryland.

Hansen, M. C., DeFries, R. S., Townshend, J. R. G., & Sohlberg, R. (2000). Global land cover classification at 1km resolution using a decision tree classifier. **International Journal of Remote Sensing**, *21*(7), 1331-1365.

Vermote, Eric; NOAA CDR Program. (2019): NOAA Climate Data Record (CDR) of AVHRR Surface Reflectance, Version 5. NOAA National Centers for Environmental Information. https://doi.org/10.7289/V53776Z4.

Vermote, Eric; NOAA CDR Program. (2022): NOAA Climate Data Record (CDR) of VIIRS Surface Reflectance, Version 1. NOAA National Centers for Environmental Information. https://doi.org/10.25921/e86y-fe34.

Vermote, E. 2018. CDR Program Climate Algorithm Theoretical Basis Document: AVHRR Land Bundle – Surface Reflectance and Normalized Difference Vegetation Index. https://www.ncei.noaa.gov/pub/data/sds/cdr/CDRs/Normalized_Difference_Vegetation_Index/AVHRR/AlgorithmDescriptionAVHRR_01B-20b.pdf

Vermote, E., J.L. Villaescusa. 2021. CDR Program Climate Algorithm Theoretical Basis Document: VIIRS Land Bundle – Surface Reflectance and Normalized Difference Vegetation Index.
https://www.ncei.noaa.gov/pub/data/sds/cdr/CDRs/Normalized_Difference_Vegetation_Index/VIIRS/AlgorithmDescriptionVIIRS_01B-20b.pdf

9. Dataset and Document Revision History

Rev 0 – 10 December 2025: This is a new document/dataset, referring to the quality-controlled monthly aggregation of Version 5 of the AVHRR Leaf Area Index CDR and Version 1 of the VIIRS Leaf Area Index CDR data set.