Leaf Area Index Climate Data Record

1. Intent of This Document and POC

1a) This document is intended for users who wish to compare satellite derived observations with climate model output in the context of the CMIP/IPCC historical experiments. Users are not expected to be experts in satellite derived Earth system observational 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 Name (as it appears on the ESGF):

--to be added once file is accepted-----

1b) Technical point of contact for this dataset:

Martin Claverie, University of Maryland, mcl@umd.edu

Eric Vermote, NASA Goddard Space Flight Center (GSFC), eric.f.vermote@nasa.gov

Jessica Matthews, NOAA National Centers for Environmental Information (NCEI) & Cooperative Institute for Climate and Satellites, North Carolina (CICS-NC), jessica.matthews@noaa.gov

2. Data Field Description

CF variable name, units:	lai, 1
Spatial resolution:	*0.05-deg in latitude and longitude*
Temporal resolution and extent:	*Monthly average, 06/1981 – 07/2016??*
Coverage:	*Global*

3. Data Origin

This dataset uses the Advanced Very High Resolution Radiometer (AVHRR) Land Bundle Surface Reflectance 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 at 2 channels and brightness temperature at 3 channels from AVHRR sensor, 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 channel 1-3. Channels 3-5 are used to retrieve atmospheric conditions, (e.g., water vapor) cloud and snow masks.

Multiple ancillary data sources are used in combination with the GAC input 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), ozone data from NASA's Total Ozone Mapping Spectrometer (TOMS), water vapor information from National Centers for Environmental Prediction (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 GAC 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.

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.

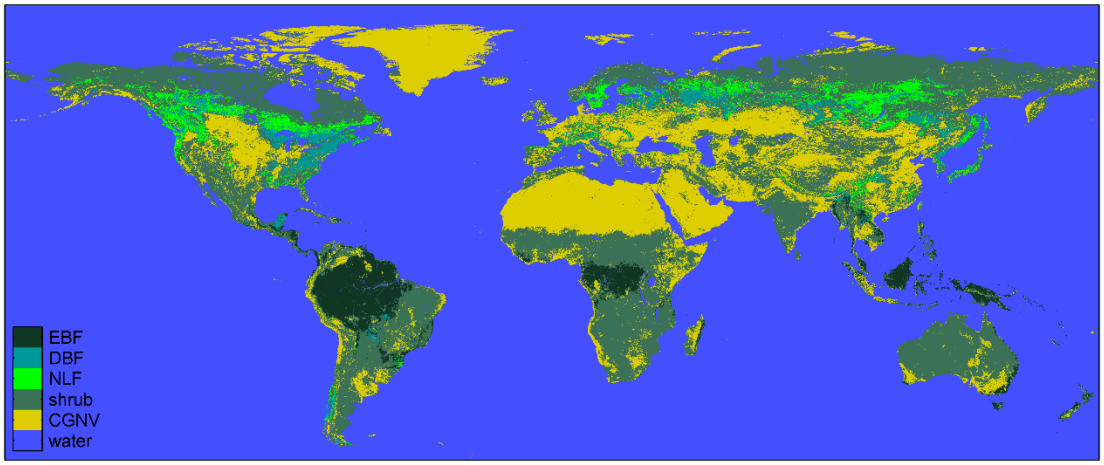

Figure 1: *Landcover classification. EBF=Evergreen broadleaf forest, DBF=Deciduous broadleaf forest, NLF=Needle leaf forest, Shrub=shrubland; CGNV=Croplands & Grasslands & Non-vegetated.*

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 2001-2007. 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).

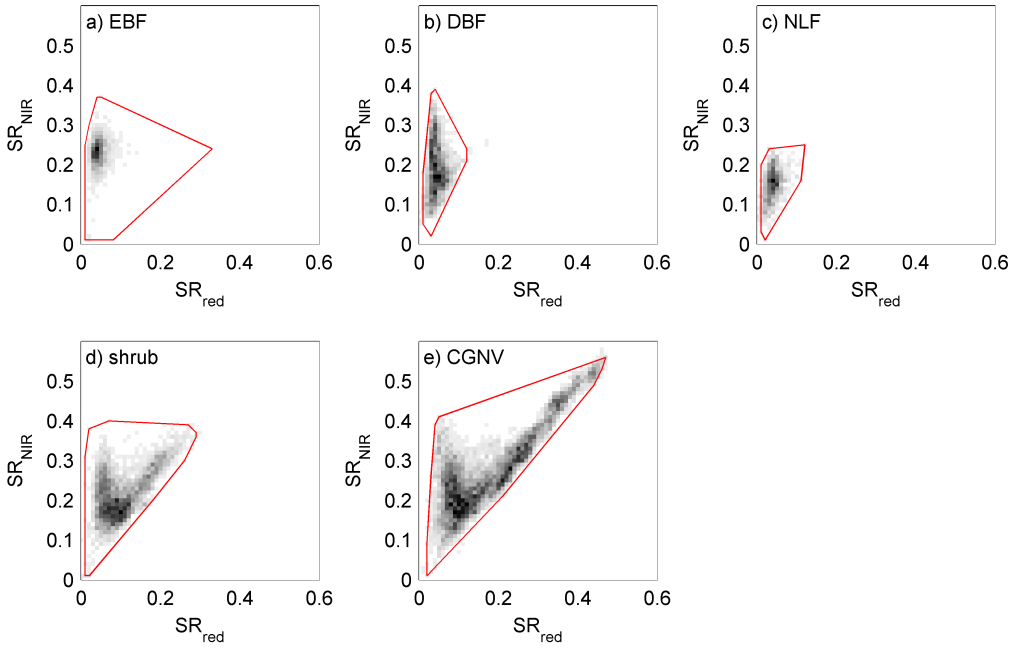

Figure 2: *Domain definition for the five classes (red polygons) in the red/NIR surface reflectance space. Greyscale images represent the density function for each 0.01 surface reflectance (SR) bin (white = no value; black = high density). Refer to Table 1 for biome class definitions. The domain definition is calculated using the AVHRR SR CDR from 2001 to 2007.*

4. Validation and Uncertainty Estimate

Direct comparison of the satellite-derived products with *in situ* measurement 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). By merging the two LAI measurements types, we obtained an uncertainty of 1.03.

Three statistical metrics (Equations (1)–(3)) are calculated: bias, the root-mean-square deviation (RMSD), and the unbiased RMSD (ubRMSD).

$$Bias = \frac{1}{n}\sum_{i=1}^{n} \varepsilon_i \quad (1)$$

$$RMSD = \sqrt{\frac{1}{n}\sum_{i=1}^{n} \varepsilon_i^2} \quad (2)$$

$$ubRMSD = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(\varepsilon_i - Bias)^2} \quad (3)$$

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 Bias, ubRMSD and RMSD from the validation over DIRECT sites. The class "Grasslands and Croplands and Non-vegetated" (CGNV) is the most represented class of the *in situ* dataset, in terms of the largest N, and shows the best result for Effective LAI. The computed RMSD fit in the medium range of previously published. 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.

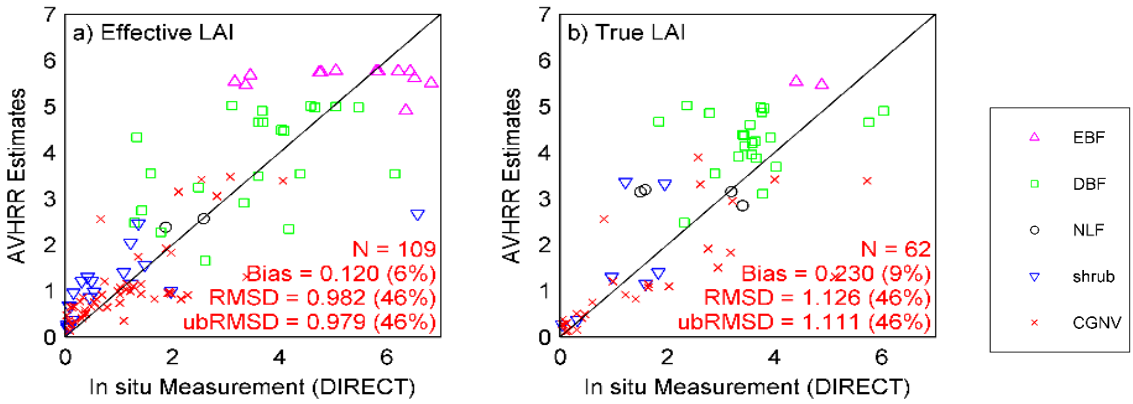

Figure 3. *In situ validation of LAI over DIRECT sites. Ground measurement covers initially a footprint of 3 km × 3 km and were extrapolated to 0.05° using MODIS products for direct comparison. Statistical metrics are defined in Equations (1)–(3); values in parenthesis correspond to metric values divided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

Class	Effective LAI				True LAI			
	Bias	**ubRMSD**	**RMSD**	**N**	**Bias**	**ubRMSD**	**RMSD**	**N**
EBF	0.24	1.33	1.31	14	0.85	0.39	0.90	2
DBF	0.36	1.27	1.29	22	0.68	1.01	1.20	22
NLF	0.25	0.37	0.36	2	0.66	1.13	1.18	4
Shrub	0.18	1.06	1.05	20	0.46	0.96	1.00	7
CGNV	−0.04	0.67	0.66	51	−0.31	1.08	1.10	27
All	**0.12**	**0.98**	**0.98**	**109**	**0.23**	**1.11**	**1.13**	**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 AVHRR Channel 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: Needle leaf Forest
		010: Broad leaf 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. Instrument Overview

Surface reflectance products are generated for each cloud-free pixel (0.05°x0.05°) observed by the AVHRR imager channel 1-3. Channels 3-5 are used to retrieve atmospheric conditions, (e.g., water vapor) cloud and snow masks. The AVHRR imager system is on board the NOAA polar-orbiting satellite series. The timeline of the NOAA platform numbers used to generate this dataset is presented in Figure 4.

Channel number	Wavelength (μm)
1	0.63
2	0.83
3a	1.61
3b	3.75
4	11.0
5	12.0

Table 3: *AVHRR channels and the associated central wavelengths. Channel 3a is only available for the AVHRR instrument on NOAA-16, -17, -18.*

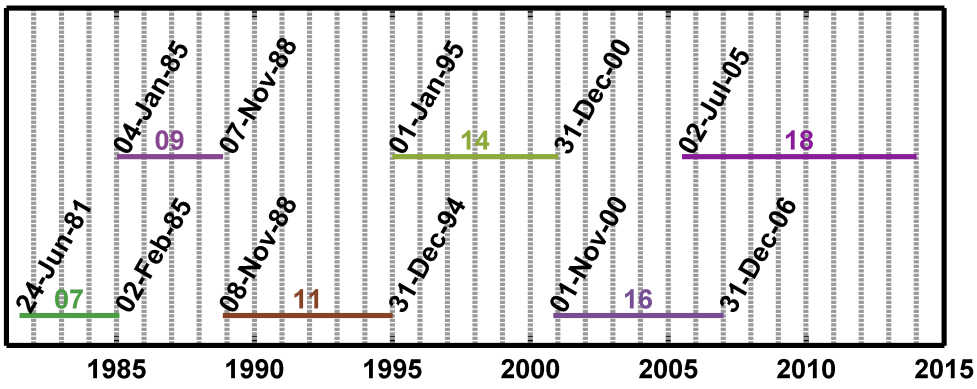

Figure 4: *Timeline of the NOAA platform numbers.*

7. References

Primary CDR references:

Claverie, M., (2014). CDR Climate Algorithm Theoretical Basis Document: Leaf Area Index (LAI) and Fraction of Absorbed Photosynthetically Active Radiation (FAPAR).

Claverie, M., E. Vermote, C. Justice, I. Csiszar, J. Eidenshink, R. Myneni, F. Baret, E. Masuoka, R. Wolfe (2014): NOAA Climate Data Record (CDR) of Leaf Area Index (LAI) and Fraction of Absorbed Photosynthetically Active Radiation (FAPAR), Version 4. NOAA National Climatic Data Center. doi: 10.7289/V5M043BX.

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, doi:10.3390/rs8030263.

Background and related references:

Baret, F.; Hagolle, O.; Geiger, B.; Bicheron, P.; Miras, B.; Huc, M.; Berthelot, B.; Nino, F.; Weiss, M.; Samain, O.; *et al.* Lai, fapar and fcover cyclopes global products derived from vegetation - part 1: Principles of the algorithm. *Remote Sens. Environ.* **2007**, *110*, 275-286.

Camacho, F.; Cemicharo, J.; Lacaze, R.; Baret, F.; Weiss, M. Geov1: LAI, FAPAR essential climate variables and FCOVER global time series capitalizing over existing products. Part 2: Validation and intercomparison with reference products. *Remote Sens. Environ.* **2013**, *137*, 310–329.

Claverie, M. 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. Toulouse 3, 2012.

Claverie, M.; Vermote, E.F.; Weiss, M.; Baret, F.; Hagolle, O.; Demarez, V. Validation of coarse spatial resolution LAI and FAPAR time series over cropland in southwest france. *Remote Sens. Environ.* **2013**, *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.; *et al.* Validation and intercomparison of global leaf area index products derived from remote sensing data. *J. Geophys. Res.-Biogeosci.* **2008**, *113*, G2.

Hansen, M., R. DeFries, J.R.G. Townshend, and R. Sohlberg (1998), UMD Global Land Cover Classification, 1 Kilometer, 1.0, Department of Geography, University of Maryland, College Park, Maryland, 1981-1994.

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

Knyazikhin, Y., J. V. Martonchik, R. B. Myneni, D. J. Diner, and S. W. Running (1998), Synergistic algorithm for estimating vegetation canopy leaf area index and fraction of absorbed photosynthetically active radiation from MODIS and MISR data, J. Geophys. Res., 103(D24), 32,257 – 32,275.

Vermote, E., C. Justice, I. Csiszar, J. Eidenshink, R. Myneni, F. Baret, E. Masuoka, R. Wolfe, M. Claverie (2014): NOAA Climate Data Record (CDR) of AVHRR Surface Reflectance, Version 4. NOAA National Climatic Data Center. doi: 10.7289/V5TM782M.

Vermote, E., and M. Claverie. 2013. CDR Climate Algorithm Theoretical Basis Document: AVHRR Land Bundle – Surface Reflectance and Normalized Difference Vegetation Index.

Vermote, E., *et al*. (2009). Towards a Generalized Approach for Correction of the BRDF Effect in MODIS Directional Reflectances. *IEEE Transactions on Geoscience and Remote Sensing*, 47, 898-908.

Vermote, E.F.; Kotchenova, S. Atmospheric correction for the monitoring of land surfaces. *J. Geophys. Res.-Atmos.* **2008**, *113*.

Weiss, M.; Baret, F.; Smith, G.J.; Jonckheere, I.; Coppin, P. Review of methods for in situ leaf area index (LAI) determination. Part ii. Estimation of LAI, errors and sampling. *Agric. For. Meteorol.* **2004**, *121*, 37–53.

Yang, W., et al. (2006), MODIS leaf area index products: From validation to algorithm improvement, IEEE Trans. Geosci. Remote Sens., 44(7), 1885 – 1898.

Zhu, Z.; Bi, J.; Pan, Y.; Ganguly, S.; Anav, A.; Xu, L.; Samanta, A.; Piao, S.; Nemani, R.; Myneni, R. Global data sets of vegetation leaf area index (LAI)3g and fraction of photosynthetically active radiation (FPAR)3g derived from global inventory modeling and mapping studies (GIMMS) normalized difference vegetation index (ndvi3g) for the period 1981 to 2011. *Remote Sens.* **2013**, *5*, 927–948.

8. Dataset and Document Revision History

Rev 0 – 29 July 2016 - This is a new document/dataset, referring to Version 4 of the Leaf Area Index CDR data set.