remote sensing

MDPI

Article

A 30+ Year AVHRR LAI and FAPAR Climate Data Record: Algorithm Description and Validation

Martin Claverie [1,2,*], Jessica L. Matthews [3,4], Eric F. Vermote [2] and Christopher O. Justice [1]

1 Department of Geographical Sciences, University of Maryland, College Park, MD 20742, USA; cjustice@umd.edu
2 NASA Goddard Space Flight Center, 8800 Greenbelt Road, Greenbelt, MD 20771, USA; eric.f.vermote@nasa.gov
3 NOAA's National Centers for Environmental Information (NCEI), 151 Patton Avenue, Asheville, NC 28801, USA; jessica.matthews@noaa.gov
4 Cooperative Institute for Climate and Satellites–NC (CICS-NC), North Carolina State University, 151 Patton Avenue, Asheville, NC 28801, USA
* Correspondence: mcl@umd.edu; Tel.: +1-301-405-4050

Academic Editors: Xuepeng Zhao, Wenze Yang, Viju John, Hui Lu, Ken Knapp, Alfredo R. Huete and Prasad S. Thenkabail
Received: 31 December 2015; Accepted: 22 February 2016; Published: 22 March 2016

Abstract: In- land surface models, which are used to evaluate the role of vegetation in the context of global climate change and variability, LAI and FAPAR play a key role, specifically with respect to the carbon and water cycles. The AVHRR-based LAI/FAPAR dataset offers daily temporal resolution, an improvement over previous products. This climate data record is based on a carefully calibrated and corrected land surface reflectance dataset to provide a high-quality, consistent time-series suitable for climate studies. It spans from mid-1981 to the present. Further, this operational dataset is available in near real-time allowing use for monitoring purposes. The algorithm relies on artificial neural networks calibrated using the MODIS LAI/FAPAR dataset. Evaluation based on cross-comparison with MODIS products and *in situ* data show the dataset is consistent and reliable with overall uncertainties of 1.03 and 0.15 for LAI and FAPAR, respectively. However, a clear saturation effect is observed in the broadleaf forest biomes with high LAI (>4.5) and FAPAR (>0.8) values.

Keywords: LAI; FAPAR; AVHRR; MODIS; climate data record; land product validation

1. Introduction

Leaf Area Index (LAI) and Fraction of Absorbed Photosynthetically Active Radiation (FAPAR) are key vegetation biophysical variables. the Global Terrestrial Observing System [1,2] has defined these variables as follow: (i) "the LAI (m^2/m^2) is geometrically defined as the total one-sided area of photosynthetic tissue per unit ground surface area"; and (ii) "the FAPAR is a primary variable controlling the photosynthetic activity of plants, and therefore constitutes an indicator of the presence and productivity of live vegetation, as well as of the intensity of the terrestrial carbon sink". These two biophysical variables are recognized by the Global Climate Observing System (GCOS) as Essential Climate Variables (ECVs, [3]). ECVs are defined as measurements of atmosphere, oceans, or land that are technically and economically feasible for systematic observation and have a high impact on the requirements of the United Nations Framework Convention on Climate Change (UNFCCC) and the Intergovernmental Panel on Climate Change (IPCC). The concept of ECVs includes a wide panel of terrestrial variables. Vegetation processes incorporated in models (*i.e.*, photosynthesis, transpiration, carbon assimilation, and respiration) are strongly driven by the surface of the plant in contact with atmosphere. In land surface models, which are used to evaluate the role of vegetation in the context of

global climate change and variability [4], LAI and FAPAR play a key role, specifically with respect to the carbon and water cycles (GCOS 2011). Hence, the GCOS and the Food and Agriculture Organization (FAO) highlighted the need for accurate LAI and FAPAR measurements from both *in situ* and remote sensing systems [1,2]. Remotely-sensed Earth observation systems provide a unique perspective to continuously monitor vegetation, globally. Using the wealth of available satellite data, the scientific community has designed ECV retrieval methods (e.g., [5–8]) and achieved an operational production of ECV global maps.

The National Oceanic and Atmospheric Administration's (NOAA) Climate Data Record (CDR) Program has the mission to support robust, sustainable, and scientifically defensible approaches to producing and preserving climate records from satellite data. The National Research Council defines a CDR as a time series of measurements of sufficient length, consistency, and continuity to determine climate variability and change [9]. The result is trustworthy information contributing to understanding how, where, and to what extent our global climate is changing. NOAA's operational CDRs are systematically generated and thoroughly documented with a six-phase research-to-operations process that includes public posting of source code used to produce the CDR.

Recently a daily surface reflectance dataset processed in a consistent and transparent way became available as an operational NOAA CDR distributed by NOAA's National Centers for Environmental Information (NCEI), available online at http://doi.org/10.7289/V5TM782M [10]. This surface reflectance dataset (named AVH09C1) has a 0.05° spatial resolution, also called Climate Modeling Grid (CMG), a daily temporal resolution, and spans from 1981 to 10 days from the present based on Advanced Very High Resolution Radiometer (AVHRR) data from seven NOAA polar orbiting satellites: NOAA-7, -9, -11, -14, -16, -18, and -19. This unique data set offers the opportunity to analyze long-term (30+ years) trends in global vegetation and as an input to generate LAI and FAPAR products. The AVHRR-derived LAI/FAPAR product is named AVH15C1.

To the author's knowledge, two other LAI/FAPAR published datasets derived from AVHRR data record are currently in existence. Both algorithms are based on machine learning using AVHRR data (surface reflectance or vegetation indices) and existing LAI/FAPAR products (e.g., the Moderate Resolution Imaging Spectroradiometer (MODIS) LAI/FAPAR product). The GLASS LAI product (http://glcf.umd.edu/data/lai/, [11]) is generated and released by the Center for Global Change Data Processing and Analysis of Beijing Normal University. The GLASS LAI product is of CMG spatial resolution (~5 km), has a temporal resolution of eight days, and is available from 1982 to 2012. However, data after 2000 are derived from the MODIS sensor rather than the AVHRR. The other existing product adopts a similar approach but relying on the AVHRR NDVI3g product (https://nex.nasa.gov/nex/projects/1349/), LAI3g and FPAR3g [12] are datasets produced by Boston University bi-monthly at the same spatial resolution, and span the period from 1981 to 2011.

The AVH15C1 LAI/FAPAR dataset will provide several advantages over both of these existing datasets. Firstly, the temporal resolution is daily, which is finer than the eight-day and 15-day resolution of GLASS and LAI3g/FPAR3g, respectively. Secondly the production of this dataset continues to present and is offered in near real-time, providing monitoring opportunities, unlike the other datasets which end in 2012 and 2011, respectively. Further, this dataset is based on the surface reflectance CDR which has been carefully geolocated, calibrated, and had necessary atmospheric and bidirectional reflectance distribution function (BRDF) corrections applied, providing a consistent AVHRR surface reflectance record as an input.

This paper provides a description of the AVH15C1 LAI/FAPAR product (called hereafter AVH15). The product is a CDR distributed by NOAA's NCEI and is available online at http://doi.org/10.7289/V5M043BX [13]. In this paper, we first list and describe all the inputs required to design, calibrate and execute the algorithm (Section 2). Section 3 is dedicated to the description of the algorithm and its calibration. Section 4 presents the evaluation and validation results.

2. Input Data

2.1. AVH09 Surface Reflectance

The AVH09C1 surface reflectance product (named hereafter AVH09) is the core input of the AVH15 CDR. This is the product distributed as the AVHRR Surface Reflectance (SR) CDR by NOAA's NCEI. It corresponds to the surface reflectance (SR) measured in two wavelengths (red, 580–680 nm, and near-infrared, 725–1000 nm) and resample at the CMG spatial resolution (0.05°). The SR is normalized for BRDF effects (nadir view and 45° solar zenith angle) using the VJB approach [14]. A full description of the product is given in the AVHRR Surface Reflectance Climate Algorithm Theoretical Basis Document (C-ATBD, [15]). The CDR distribution of AVH09 contains associated quality information such as indications of cloud, cloud-shadow, and snow-free conditions. NDVI (Normalized Difference Vegetation Index) was derived directly from the nadir-adjusted surface reflectance.

While the AVH09 product is a measurement of SR normalized to a constant 45° solar zenith angle, the FAPAR varies according to solar zenith angle. To be consistent with the FAPAR definition used in comparable FAPAR products, such as the MODIS product (which is an instantaneous FAPAR at the time of the satellite overpass), FAPAR is derived from nadir-adjusted surface reflectance with the solar zenith angle at the time of the acquisition. The same VJB BRDF model performs this adjustment.

The VJB model is based on two coefficients (V and R), describing the 3D structure of the surface, that have been demonstrated to be well correlated to NDVI [14]. Based on MODIS CMG data, the linear regression coefficients were retrieved globally over the land surface. The retrieval of the nadir-adjusted and variable solar zenith (θ_s) angle SR ($\rho\left(O_S, 0, 0\right)$) from the AVH09 product is computed as in Equation (1), where $\rho^N\left(45, 0, 0\right)$ is the AVH09 SR and V and R are the two coefficients of the VJB model. F1 and F2 are two kernels defined by [16] and [17], respectively, that both depend on illumination-view geometry.

$$\rho\left(O_S, 0, 0\right) = \rho^N\left(45, 0, 0\right) \frac{1 + V \times F_1\left(O_S, 0, 0\right) + R \times F_2\left(O_S, 0, 0\right)}{1 + V \times F_1\left(45, 0, 0\right) + R \times F_2\left(45, 0, 0\right)} \quad (1)$$

2.2. Reference LAI/FAPAR

The MCD15A2 (denoted MCD15 hereafter) is used as the reference LAI/FAPAR dataset. It is a global LAI and FAPAR product. MCD15 is used as a LAI/FAPAR reference to calibrate the AVH15 algorithm. MCD15 is composited every eight days at 1-km resolution on a sinusoidal grid. It is produced on near real-time during the entire MODIS era (2000 to present). The main algorithm is based on lookup tables (LUT) simulated from a 3D radiative transfer model [18]. Red and NIR (Near infrared) atmospherically-corrected MODIS reflectance [19] and the corresponding illumination-view geometry are used as inputs of the LUT. The output is the mean LAI and FAPAR values computed over the set of acceptable LUT elements for which simulated and MODIS surface reflectances agree within specified level of (model and measurement) uncertainties. When the main algorithm fails, a backup algorithm based on NDVI relationships, calibrated over the same radiative transfer model simulations is used [20]. Retrievals from both algorithms were used. LAI and FAPAR are first produced daily. Then, the eight-day composite corresponds to the values of the product when the maximum FAPAR value within the eight day period is observed. Note that no LAI and FAPAR values are retrieved over bare or very sparsely vegetated area, permanent ice or snow area, permanent wetland, urban area, or water bodies.

2.3. Land Cover Map

Land cover classification, used to stratify the outputs, follows the 1981–1994 IGBP (International Geosphere-Biosphere Program) map [21]. To avoid inconsistent land cover changes, the same classification is used for the entire dataset. Moreover, to reduce spatial discontinuity, the number of classes was reduced from 17 to 6 (see Table 1). A global map depicting the land cover categories used is shown in Figure 1.

Table 1. Reclassification table of IGBP land cover classes.

Short Name	New Class Name	Original IGBP Class Name
NLF	Needle leaf forest	Needle leaf forest
DBF	Deciduous broadleaf forest	Deciduous broadleaf forest, mixed forests
Shrub	Shrubland	Closed/open/woody shrubland, savannas
CGNV	Croplands & grasslands & non-vegetated	Grasslands, permanent wetlands, croplands, urban and built-up, cropland/natural vegetation mosaic, snow and ice, barren or sparsely vegetated
EBF	Evergreen broadleaf forest	Evergreen broadleaf forest
water	Water	Water

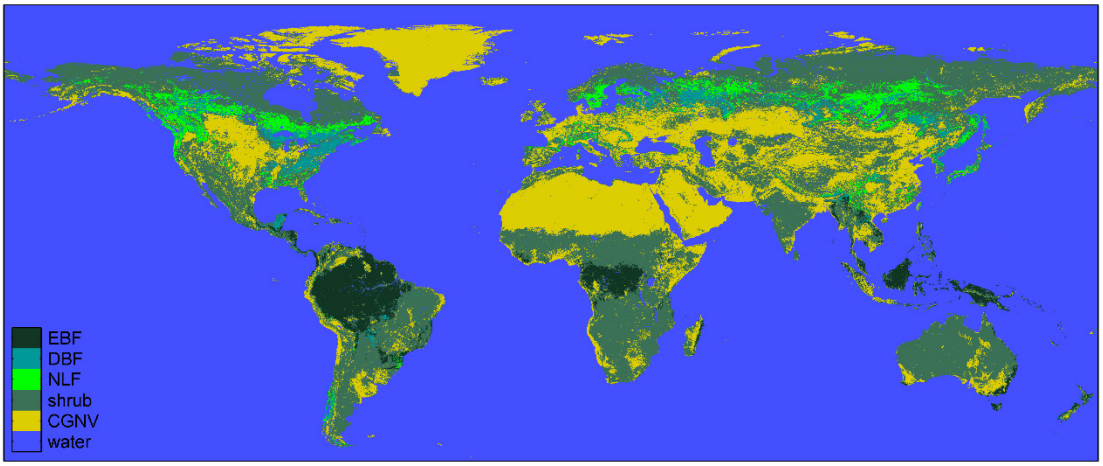

Figure 1. International Geosphere-Biosphere Program (IGBP) land cover classification from Hansen *et al*. [21]. Labels were simplified according to Table 1.

2.4. Calibration and Validation Sites

For calibration and validation purposes, a set of globally-distributed sites were selected. Two networks covering most land biomes were used: the BELMANIP2 network for calibration and the DIRECT network for validation. Figure 2 illustrates the global coverage and distribution of sites from both networks.

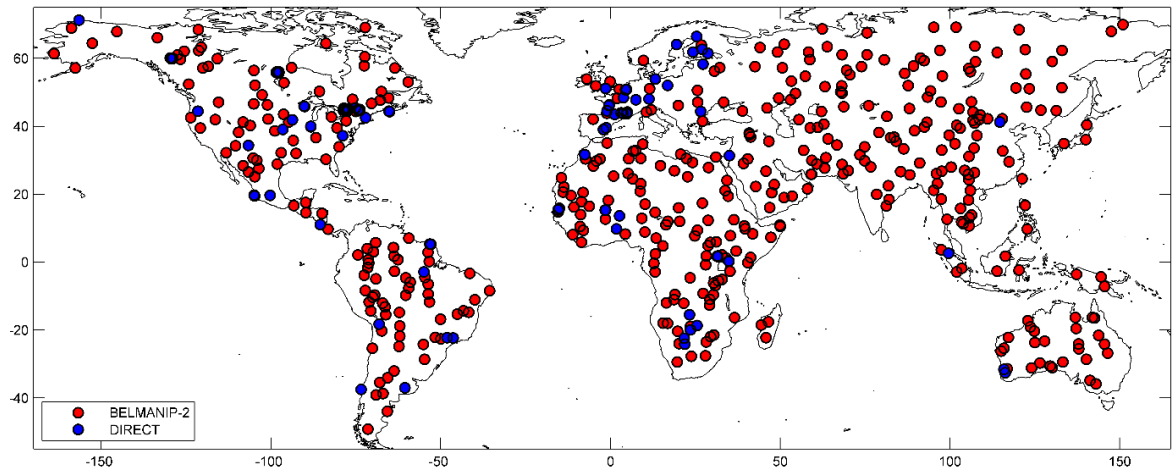

Figure 2. BELMANIP-2 and DIRECT network sites location.

The BELMANIP2 (BEnchmark Land Multisite ANalysis and Intercomparison of Products 2, updated version of BELMANIP1, [22]) network was created using sites from existing experimental networks (FLUXNET, AERONET, VALERI, BigFoot, ...) completed with selected sites from the GLC2000 land cover map. The site selection was performed for each band of latitude (10° width) by keeping the same proportion of biome types within the selected sites as within the whole band of latitude. Attention was paid so that the sites were homogeneous over a 10×10 km^2 area, almost flat, and with a minimum proportion of urban area and permanent water bodies. The BELMANIP2 dataset includes 445 sites.

DIRECT is a collection of sites for which ground measurements have been collected [23] and processed according to the CEOS-LPV (Centre for Earth Observation Science—Land Product Validation) guidelines. This network is used for evaluation purposes since it contains sites different from those in BELMANIP2. There are currently 113 *in situ* data points available. Data were obtained through the On-Line Validation Exercise (OLIVE) a CEOS-LPV initiative for online validation of global land products [24].

Matchups with MCD15 and AVH09 were performed by selecting pixels whose center is the nearest to the BELMANIP2 and DIRECT site locations. We focused on coincident years between NOAA-16 and MODIS products: 2001–2007. Since MCD15 products are 1 km resolution, they were aggregated to 0.05° to be consistent with the spatial resolution of AVH09. Finally, for temporal consistency and to reduce the noise, AVH09 and MCD15 products were averaged monthly.

3. Algorithm Definition and Calibration

The AVH15 algorithm is based on an artificial neural network (ANN) connecting LAI or FAPAR to AVH09 SR products for each of the five biomes (omitting the water class) as defined in Section 2.3. The algorithm has five steps:

- Input normalization,
- ANN execution (per class and variable),
- Output normalization,
- Classes fusion according to the IGBP land cover as defined in Section 2.3, and
- Flagging pixels outside of the defined domain.

The ANN were trained using the MCD15 products as reference. The calibration procedure was done using data from BELMANIP2 as described in Section 2.4. We retained data for the period spanning from 2001 to 2007.

3.1. Input and Output Normalization

Normalization is achieved simply by scaling between the minimum and maximum values: the normalized values (Y) will vary between −1 and +1, and are computed from the raw values (X), and the minimum (X_{min}) and maximum (X_{max}) values (Equation (2)). Minimum and maximum values of input and outputs retrieved with the calibration data are reported per class (*i.e.*, for each ANN) in Table 2.

$$Y = \frac{2 \times (X - Xmin)}{(Xmax - Xmin)} - 1 \quad (2)$$

3.2. ANN Definition and Learning

ANN are connections of "neurons", associated by "synaptic" weights. Each neuron transforms the sum of the weighted signal from the previous neurons according to a given transfer function and a bias. The combination of sigmoid and linear functions is capable to fit any type of function [25].

Table 2. Minimum and maximum values used for normalization. Refer to Table 1 for biome class definition.

Class	EBF		DBF		NLF		Shrub		CGNV	
	Min	Max	Min	Max	Min	Max	Min	Max	Min	Max
Red	0.01	0.31	0.01	0.12	0.01	0.11	0.01	0.29	0.02	0.42
NIR	0.01	0.37	0.04	0.39	0.02	0.24	0.01	0.39	0.02	0.48
$\cos(\theta_s)$	0.46	0.88	0.14	0.88	0.05	0.85	0.06	0.88	0.01	0.88
NDVI	−0.41	0.91	−0.01	0.87	0.01	0.86	−0.22	0.92	−0.22	0.80
LAI	0.69	6.72	0.01	5.94	0.00	5.24	0.00	5.95	0.00	5.27
FAPAR	0.23	0.91	0.02	0.92	0.01	0.93	0.00	0.89	0.00	0.89

The ANN architecture finally retained followed the proposition of [6] to have as many intermediate neurons as inputs plus one. The ANN architecture is composed of (see Figure 3):

- One input layer made of the four normalized inputs: AVH09 Red SR, AVH09 NIR SR, AVH13 NDVI and the cosine of the solar zenith angle,
- One hidden layer with five neurons having hyperbolic tangent sigmoid transfer functions (Equation (3), where x is the neuron input and y the output),
- One output layer via a linear transfer function, and
- Normalized output.

$$y = \frac{2}{1 + \exp(-2 * x)} - 1 \tag{3}$$

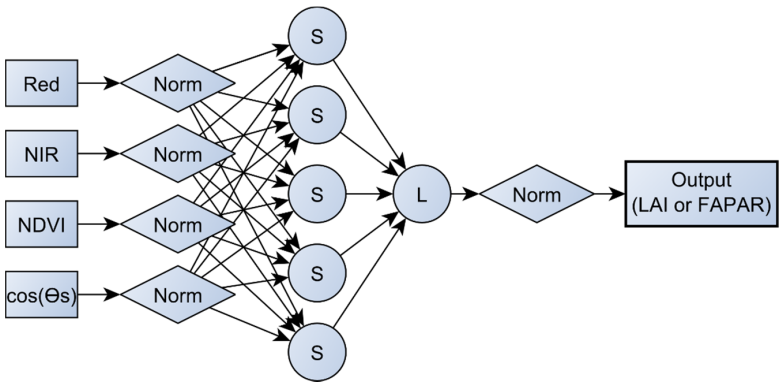

Figure 3. Conceptual representation of the ANN, including normalization ("Norm") steps. Notice that the number of neurons correspond to the actual number of neurons. S and L stand for "sigmoid" and "linear" neurons, respectively.

For each of the 10 configurations (*i.e.*, five classes × two output variables), 10 ANN were trained, resulting to 100 ANN in total. The selection of the 10 optimal ANN was based on the RMSD (root-mean-square deviation) between the outputs and the *in situ* data from DIRECT network sites.

The learning process is composed of two elements: the training dataset that was described previously, and the learning rule that is now described. The Levenberg–Marquardt optimization algorithm [26] is used to adjust the synaptic weights and neuron bias to get the best agreement between the output simulated by the network and the corresponding value of MCD15 LAI/FAPAR. The initial values of the weights and biases were set to a random value between −1 and +1.

3.3. Domain Definition

The ANN are trained over a defined area and the output accuracy decreases considerably outside of the domain delimited by the learning dataset. Therefore, we defined an acceptable input domain for

each class based on SR inputs used during ANN training. Since output data are not considered, no differentiation is necessary between LAI and FAPAR domain definitions.

The domain definition was defined based on calibration data from NOAA-16 AVHRR SR pixels overlapping spatially and temporally with BELMANIP2 sites during the 2001–2007 period. Figure 4 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). With these selected pixels, a convex hull algorithm is used to define the exact envelope. Finally, the envelope was simplified using a recursive Douglas–Peucker polyline simplification algorithm [27].

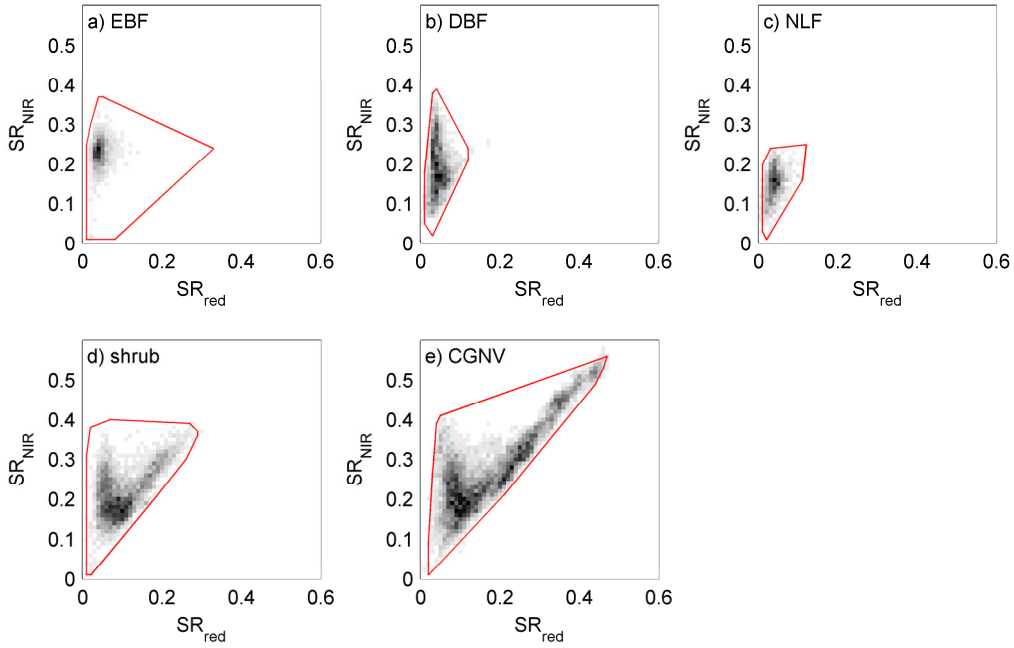

Figure 4. (**a**–**e**) Domain definition for the five classes (red polygons) in the red/NIR surface reflectance space. Greyscale images represent the density function for each 0.01 surface reflectance (SR) bin (white = no value; black = high density). Refer to Table 1 for biome class definitions. The domain definition is calculated using AVH09C1 data acquired from 2001 to 2007.

Each processed pixel is compared to the polygon of the corresponding class. If it falls outside the polygon, an indicator flag is reported in the QA associated with the retrieval.

4. CDR Performance and Validation

CDR performance is evaluated by examining the theoretical performance of the ANN as presented in Figure 5. Three statistical metrics (Equations (4)–(6), [28]) are calculated: bias, the root-mean-square deviation (RMSD), and the unbiased RMSD (ubRMSD).

$$Bias = \frac{1}{n} \times \sum_{i=1}^{n} \varepsilon_i \tag{4}$$

$$RMSD = \sqrt{\frac{1}{n} \times \sum_{i=1}^{n} {\varepsilon_i}^2} \tag{5}$$

$$ubRMSD = \sqrt{\frac{1}{n-1} \times \sum_{i=1}^{n} (\varepsilon_i - Bias)^2} \tag{6}$$

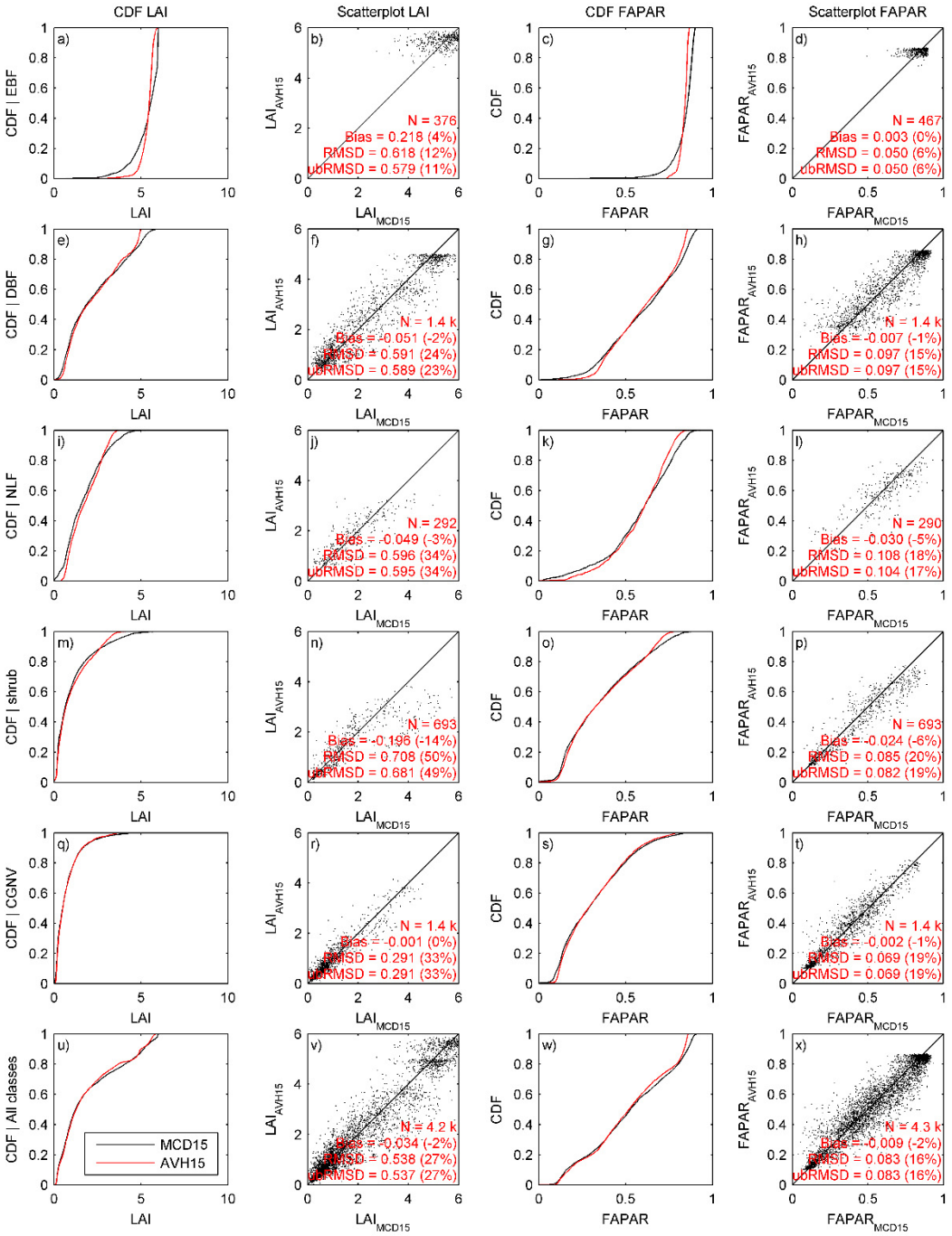

Figure 5. (**a**–**x**) Theoretical performances of the LAI and FAPAR retrieval. Each row refers to a land cover class and the bottom row to all classes merged. On the first column, the cumulative distribution functions (CDF) of MCD15 training data and AVH15 LAI retrieval are shown. On the second column, scatter plots between LAI MCD15 (x-axis) and LAI AVH15 (y-axis) are displayed. The graphs are reproduced for FAPAR on the third and fourth column. Only data from DIRECT sites (not used for training) were plotted. Statistical metrics on the second and fourth subplot columns are defined in Equations (4)–(6); values in parenthesis correspond to metric values divided by the reference mean value. Refer to Table 1 for class biome definitions.

In Equations (4)–(6), 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 data shown in Figure 5 are based on DIRECT site locations, which were not used for the training but only for the evaluation of the ANN. The cumulative distribution functions demonstrate the good overall performance of the ANN to reproduce the range of value. However, a clear saturation effect is observed in the DBF and EBF classes with high LAI (>4.5) and FAPAR (>0.8) values. The good performance of non-broadleaf-forest classes (*i.e.*, NLF, shrub, and CGNV) is consistent with the fact that the input data used for calibration (*i.e.*, MCD15) have shown better performance on such land cover [6,29].

Another way to evaluate the reproducibility of the algorithm is to compare outputs from AVHRR sensors on board two different platforms. We selected NOAA-16 and NOAA-18, with an overlapping period from 2 July 2005 to 31 December 2006. The analysis is carried on BELMANIP2 and DIRECT sites. The resultant scatterplot is displayed in Figure 6. RMSD are 0.35 for LAI and 0.07 for FAPAR, which are comparatively better than the ANN performances computed in Figure 5l,x comparing AVHRR-derived to MODIS-derived LAI and FAPAR, respectively.

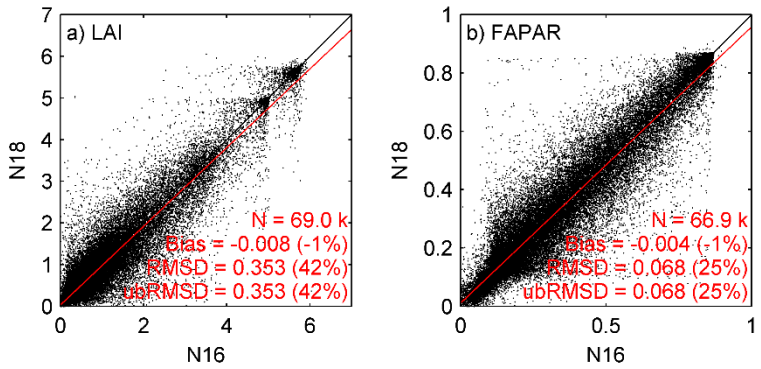

Figure 6. (**a**,**b**) Comparison of retrieval from AVHRR NOAA-16 (N16) and AVHRR NOAA-18 (N18) for BELMANIP-2 and DIRECT sites from 2 July 2005 to 31 December 2006. Statistical metrics are defined in Equations (4)–(6); values in parenthesis correspond to metric values divided by the reference mean value.

Direct comparison of the satellite-derived products with *in situ* measurement is a key validation step. It has been defined as Stage 1 of the CEOS-LPV guideline. 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 [30]. We relied on the work performed by Garrigues *et al*. [23] who contributed to the conception of the DIRECT network. They first gathered *in situ* measurements from many locations and scaled them up to a 3 km × 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 LAI/FAPAR MCD15 (1 km) retrieval aggregated over the measurement footprint (3 km × 3 km) and the one aggregated at 0.05°. The outputs were finally compared to the AVH15 retrieval (Figure 7). Notice that these DIRECT sites are independent of the ANN learning process. LAI validation scatter plots were separated among the type of measurement: effective and true LAI, depending if the clumping factor is considered or not [6,31]. By merging the two LAI measurements types, we obtained an uncertainty of 1.03.

The error budget is detailed in Table 3, 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*. [32] validated four global LAI/FAPAR products (GEOV1, CYCLOPE, MCD15, and GLOV2) and found RMSD ranges of (0.74–1.39) and (0.078–0.228) for effective LAI and FAPAR, respectively.

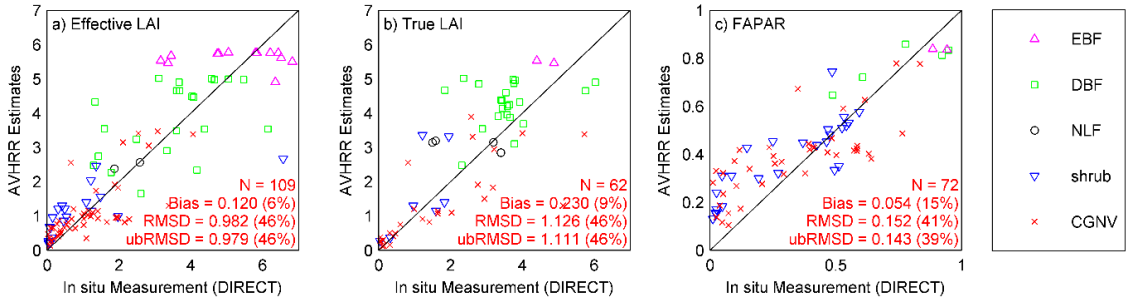

Figure 7. (**a**–**c**) *In situ* validation over DIRECT sites. Ground measurement covers initially a footprint of 3 km × 3 km and were extrapolated to 0.05° using MCD15 products for direct comparison. Statistical metrics are defined in Equations (4)–(6); values in parenthesis correspond to metric values divided by the reference mean value.

Table 3. Error budget based on in situ validation. N corresponds to the number of points used to compute the statistical metrics.

Class	Effective LAI				True LAI				FAPAR			
	Bias	**ubRMSD**	**RMSD**	**N**	**Bias**	**ubRMSD**	**RMSD**	**N**	**Bias**	**ubRMSD**	**RMSD**	**N**
EBF	0.24	1.33	1.31	14	0.85	0.39	0.90	2	−0.08	0.04	0.08	2
DBF	0.36	1.27	1.29	22	0.68	1.01	1.20	22	0.03	0.13	0.12	5
NLF	0.25	0.37	0.36	2	0.66	1.13	1.18	4	N/A	N/A	N/A	0
Shrub	0.18	1.06	1.05	20	0.46	0.96	1.00	7	0.08	0.12	0.14	25
CGNV	−0.04	0.67	0.66	51	−0.31	1.08	1.10	27	0.05	0.16	0.16	40
All	**0.12**	**0.98**	**0.98**	**109**	**0.23**	**1.11**	**1.13**	**62**	**0.05**	**0.14**	**0.15**	**72**

5. Conclusions

This paper presented the 30+ year AVH15 LAI/FAPAR CDR distributed by the NOAA's NCEI. The dataset is a global, daily, 0.05° (~5 km) spatial resolution, spanning from 1982 to 10 days from present. The algorithm relies on artificial neural networks which were calibrated per land cover type using the MODIS LAI/FAPAR product. Five land cover classes were included: evergreen broadleaf forest, deciduous broadleaf forest, needle leaf forest, shrubland, croplands and grasslands, and non-vegetated.

Reproducibility of the algorithm was demonstrated to achieve overall uncertainties performance of 0.54 (27%) and 0.08 (16%), for LAI and FAPAR, respectively. However, per-biome scores are contrasted. Best performances were computed for cropland, grassland, and non-vegetated surfaces. The main limitation of the algorithm is the incapacity to reproduce variability of densely vegetated cover such as deciduous forest. A clear saturation of the algorithm was observed in these biomes for high LAI (>4.5) and FAPAR (>0.8) values. The outputs were also compared to *in situ* data. Overall uncertainties of 0.98 and 1.13 were found for LAI, depending on the accounting of the clumping index in the *in situ* processing. Overall uncertainties of 0.15 were found for FAPAR. The high values of these scores are also related to the complexity of producing a consistent measurement over a pixel footprint (*i.e.*, 0.05° × 0.05°).

Acknowledgments: The authors would like to acknowledge NOAA NCEI for supporting the development of the AVH15 product. J.L.M. was supported by NOAA through the Cooperative Institute for Climate and Satellites - North Carolina under Cooperative Agreement NA14NES432003.

Author Contributions: Martin Claverie developed the AVH15C1 algorithm, conducted the data analysis and wrote the paper. Eric F. Vermote supervised the work and contributed with ideas. Jessica L. Matthews and Christopher O. Justice reviewed this manuscript and provided comments and suggestions.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Gobron, N.; Verstraete, M. *Ecv t10 Fapar*; GTOS: Rome, Italy, 2009.
2. Gobron, N.; Verstraete, M. *Ecv t11 Leaf Area Index (LAI)*; GTOS: Rome, Italy, 2009.
3. GCOS. *Systematic Observation Requirements for Satellite-Based Products for Climate (2011 Update)*; GCOS-154; GCOS: Geneva, Switzerland, 2011.
4. Running, S.W.; Baldocchi, D.D.; Turner, D.P.; Gower, S.T.; Bakwin, P.S.; Hibbard, K.A. A global terrestrial monitoring network integrating tower fluxes, flask sampling, ecosystem modeling and EOS satellite data. *Remote Sens. Environ.* **1999**, *70*, 108–127. [CrossRef]
5. 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. [CrossRef]
6. 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. [CrossRef]
7. Myneni, R.B.; Hoffman, S.; Knyazikhin, Y.; Privette, J.L.; Glassy, J.; Tian, Y.; Wang, Y.; Song, X.; Zhang, Y.; Smith, G.R.; *et al*. Global products of vegetation leaf area and fraction absorbed par from year one of MODIS data. *Remote Sens. Environ.* **2002**, *83*, 214–231. [CrossRef]
8. Weiss, M.; Baret, F.; Garrigues, S.; Lacaze, R. Lai and fapar cyclopes global products derived from vegetation. Part 2: Validation and comparison with MODIS collection 4 products. *Remote Sens. Environ.* **2007**, *110*, 317–331.
9. National Research Council. *Climate Data Records from Environmental Satellites: Interim Report*; The National Academies Press: Washington, DC, USA, 2004; p. 150.
10. Vermote, E.; Justice, C.; Csiszar, I.; Eidenshink, J.; Myneni, R.; Baret, F.; Masuoka, E.; Wolfe, R.; Claverie, M.; Program, N.C. *Noaa Climate Data Record (CDR) of Avhrr Surface Reflectance*; version 4; NOAA National Climatic Data Center, Ed.; NOAA National Climatic Data Center: Asheville, NC, USA, 2014.
11. Xiao, Z.; Liang, S.; Wang, J.; Chen, P.; Yin, X.; Zhang, L.; Song, J. Use of general regression neural networks for generating the glass leaf area index product from time-series MODIS surface reflectance. *IEEE Trans. Geosci. Remote Sens.* **2014**, *52*, 209–223. [CrossRef]
12. 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.
13. Claverie, M.; Vermote, E.; Program, N.C. *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, Ed.; NOAA National Climatic Data Center: Asheville, NC, USA, 2014.
14. Vermote, E.; Justice, C.O.; Breon, F.-M. Towards a generalized approach for correction of the BRDF effect in MODIS directional reflectances. *IEEE Trans. Geosci. Remote Sens.* **2009**, *47*. [CrossRef]
15. Vermote, E.; Claverie, M. *Avhrr Land Bundle - Surface Reflectance and Normalized Difference Vegetation Index: Climate Algorithm Theoretical Basis Document*; NOAA National Climatic Data Center: Asheville, NC, USA, 2013.
16. Roujean, J.-L.; Leroy, M.; Deschamps, P.Y. A bidirectional reflectance model of the earth's surface for the correction of remote sensig data. *J. Geophys. Res.* **1992**, *97*, 20455–20468. [CrossRef]
17. Li, X.W.; Strahler, A.H. Geometric-optical bidirectional reflectance modeling of the discrete crown vegetation canopy: Effect of crown shape and mutual shadowing. *IEEE Trans. Geosci. Remote Sens.* **1992**, *30*, 276–292. [CrossRef]
18. Knyazikhin, Y.; Martonchik, J.V.; Myneni, R.B.; Diner, D.J.; Running, S.W. Synergistic algoritm for estimating vegetation canopy leaf area index and fraction of absorbed photosynthetically active radiation from MODIS and MISR data. *Geophys. Res.* **1998**, *103*, 32257–32276. [CrossRef]
19. Vermote, E.F.; Tanre, D.; Deuze, J.L.; Herman, M.; Morcrette, J.J. Second simulation of the satellite signal in the solar spectrum, 6S: An overview. *IEEE Trans. Geosci. Remote Sens.* **1997**, *35*, 675–686. [CrossRef]

20. Yang, W.; Tan, B.; Huang, D.; Rautiainen, M.; Shabanov, N.; Wang, Y.; Privette, J.; Huemmrich, K.; Fensholt, R.; Sandholt, I.; *et al*. MODIS leaf area index products: From validation to algorithm improvement. *IEEE Trans. Geosci. Remote Sens.* **2006**, *44*, 1885–1898. [CrossRef]
21. Hansen, M.; DeFries, R.; Townshend, J.R.G.; Sohlberg, R. *Umd Global Land Cover Classification, 1 Kilometer, 1981–1994, 1.0*; Department of Geography, University of Maryland: College Park, MD, USA, 1998.
22. Baret, F.; Morissette, J.; Fernandes, R.; Champeaux, J.; Myneni, R.; Chen, J.; Plummer, S.; Weiss, M.; Bacour, C.; Garrigues, S.; *et al*. Evaluation of the representativeness of networks of sites for the global validation and intercomparison of land biophysical products: Proposition of the CEOS-BELMANIP. *IEEE Trans. Geosci. Remote Sens.* **2006**, *44*, 1794–1803. [CrossRef]
23. 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. [CrossRef]
24. Weiss, M.; Baret, F.; Block, T.; Koetz, B.; Burini, A.; Scholze, B.; Lecharpentier, P.; Brockmann, C.; Fernandes, R.; Plummer, S.; *et al*. On line validation exercise (olive): A web based service for the validation of medium resolution land products. Application to fapar products. *Remote Sens.* **2014**, *6*, 4190–4216. [CrossRef]
25. Demuth, H.; Beale, M. *Neural Network Toolbox for Use with Matlab*; The MathWorks Inc.: Natick, MA, USA, 1998.
26. Moré, J. The levenberg-marquardt algorithm: Implementation and theory. In *Numerical Analysis*; Watson, G.A., Ed.; Springer: Berlin, Germany; Heidelberg, Germany, 1978; Volume 630, pp. 105–116.
27. Douglas, D.H.; Peucker, T.K. Algorithms for the reduction of the number of points required to represent a digitized line or its caricature. In *Classics in Cartography*; John Wiley & Sons, Ltd.: Hoboken, NJ, USA, 2011; pp. 15–28.
28. Vermote, E.F.; Kotchenova, S. Atmospheric correction for the monitoring of land surfaces. *J. Geophys. Res.-Atmos.* **2008**, *113*, D23. [CrossRef]
29. 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. [CrossRef]
30. Roman, M.O.; Gatebe, C.K.; Shuai, Y.; Wang, Z.; Gao, F.; Masek, J.G.; He, T.; Liang, S.; Schaaf, C.B. Use of in situ and airborne multiangle data to assess MODIS-and landsat-based estimates of directional reflectance and albedo. *IEEE Trans. Geosci. Remote Sens.* **2013**, *51*, 1393–1404. [CrossRef]
31. 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.
32. 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. [CrossRef]