DSCOVR EPIC Vegetation Earth System Data Record

Science Data Product Guide

Version 2

Yuri Knyazikhin and Ranga B. Myneni

Boston University

Boston, January 10, 2021

TABLE OF CONTENTS

1. INTRODUCTION **2**

2. EXPERIMENT OVERVIEW **4**

3. SUN-SENSOR GEOMETRY **5**

5. LEVEL 2 VESDR PRODUCT **8**

5.1. VESDR product file name *8*
5.2. HDF file structure *8*
5.3. Root and tile attributes *9*
5.4. Datasets *10*
5.5. Quality assessment dataset *11*

6. ANCILLARY SCIENCE DATA PRODUCTS **14**

6.1. DSCOVR EPIC land cover type *14*
6.2. Distribution of land cover types. *17*

7. KNOWN ISSUES **19**

8. EXAMPLES **19**

1. INTRODUCTION

1.1. Purpose. This document describes Version 2 Level 2 Vegetation Earth System Data Record (VESDR) derived from the Earth Polychromatic Imaging Camera (EPIC) onboard NOAA's Deep Space Climate Observatory (DSCOVR). It provides file structure for the geophysical and ancillary science data products. The VESDR parameters are summarized in Table 1.

Table 1: Vegetation Parameter Suite in the Level 2 Vegetation Earth System Data Record (VESDR)

Parameter name	Units	Resolution		Comments
		Temporal	Spatial	
Normalized Difference Vegetation Index (NDVI)	none	65 to 110 min	10018.7542 m	Difference between Reflectance Factor (BRF) at 779.5 nm and 680 nm normalized by their sum
Fraction vegetation absorbed Photosynthetically Active Radiation (FPAR)	fraction	65 to 110 min	10018.7542 m	Fraction of photosynthetically active radiation (400 – 700nm) absorbed by vegetation
Leaf Area Index (LAI)	$\frac{m^2_{plant}}{m^2_{ground}}$	65 to 110 min	10018.7542 m	One-sided green leaf area per unit ground area in broadleaf canopies and the projected needle area in coniferous canopies
Sunlit Leaf Area Index (SLAI)	$\frac{m^2_{sunlit}}{m^2_{ground}}$	65 to 110 min	10018.7542 m	Sunlit green leaf area per unit ground area
Precision of Leaf Area Index (Dlai)	$\frac{m^2_{plant}}{m^2_{gound}}$	65 to 110 min	10018.7542 m	Retrieval dispersion of LAI
Directional Area Scattering Factor (DASF)	none	65 to 110 min	10018.7542 m	Estimate of Canopy Bidirectional Reflectance Factor as if the foliage does not absorb radiation
Earth Reflector Type Index (ERTI)	none	65 to 110 min	10018.7542 m	Estimate of the recollision probability p transformed to the interval [0°,180°] as $\mathrm{atan}(p)$ if $\mathrm{atan}\, p \geq 0$ and $\mathrm{atan}(p) + 180°$ otherwise.
Scattering coefficient at 443 nm	none	65 to 110 min	10018.7542 m	Estimate of the fraction of intercepted radiation that has been reflected from, or diffusively transmitted through, the vegetation at 443 nm.
Scattering coefficient at 551 nm	none	65 to 110 min	10018.7542 m	... at 551 nm
Scattering coefficient at 680 nm	none	65 to 110 min	10018.7542 m	... at 680 nm
Scattering coefficient at 780 nm	none	65 to 110 min	10018.7542 m	... at 780 nm
Quality Assessment variable	none	65 to 110 min	10018.7542 m	Overall quality of the VESDR parameters and 'Status_QA' copied from DSCOVR EPIC L2 MAIAC (version 2)
Aerosol Optical Depth at 443 nm	none	65 to 110 min	10018.7542 m	AOD at 443 nm copied from upstream DSCOVR EPIC L2 MAIAC (version 2) product
Aerosol Optical Depth at 551 nm	none	65 to 110 min	10018.7542 m	AOD at 443 nm copied from upstream DSCOVR EPIC L2 MAIAC (version 2) product
Cloud Mask and Land- Water Mask	none	65 to 110 min	10018.7542 m	Cloud mask and Land-Water mask copied from upstream DSCOVR EPIC L2 MAIAC (version 2) product

The VESDR file also includes Solar Zenith Angle (SZA), Solar Azimuthal Angle (SAA), View Zenith (VZA) and Azimuthal (VAA) angles at the same temporal and spatial resolutions (Sect. 3).

The DSCOVR EPIC Science Algorithm Team also provides two ancillary science data products, namely, 10018.7542m Land Cover Type and Distribution of Land Cover Types within 10 km EPIC pixel. The products were derived from 500m MODIS land cover type 3 product (MCDLCHKM), which was generated from 2008, 2009 and 2010 land cover products (MCD12Q1, v051). The ancillary data sets are summarized in Table 2.

Table 2: Ancillary science data product derived form 500m MODIS land cover type 3 product

Parameter name	Units	Resolution		Comments
		Temporal	Spatial	
Land Cover Type	none	static	10018.7542m	Regional SIN Land Cover Type maps
Land Cover Type Distribution	none	static	10018.7542m	Distribution of land cover types within 10 km EPIC pixel

All products are projected on four regional 10018.7542m sinusoidal (SIN) grids (Sect. 4) and written in the standard Hierarchical Data Format 5 (HDF5) using HDF-defined data models (http://www.hdfgroup.org/HDF5/). The EPIC VESDR and ancillary data products are publicly available from the NASA Langley Atmospheric Science Data Center.

1.2. Product maturity level. Definitions of product maturity levels developed by the MISR team are adopted (https://www-misr.jpl.nasa.gov/getData/maturityLevels/). The DSCOVR EPIC version 2 VESDR product is released at Provisional quality level, i.e.,

- Incremental improvements are still occurring. Obvious artifacts or blunders observed in prerelease product have been identified and either minimized or documented
- General research community is encouraged to participate in the quality assessment and validation, but need to be aware that product validation and quality assessment are ongoing
- Parameter may be used in publications as long as provisional quality is indicated by the authors. Users are urged to contact science team representatives prior to use of the data in publications, and to recommend members of the instrument teams as reviewers
- The Data Quality Summary states estimated uncertainties
- May be replaced in the archive when an upgraded product becomes available, but should be reproducible upon demand

DSCOVR EPIC data products begin in a provisional state, and advance through a series of maturity levels, from Provisional to Validated status, i.e. from a developmental status to a scientifically proven status.

1.3. DSCOVR EPIC documents. Project documents are available at https://eosweb.larc.nasa.gov/project/DSCOVR/DSCOVR_EPIC_L2_VESDR_02. DSCOVR EPIC publications can be found at https://epic.gsfc.nasa.gov/science/pubs. The VESDR theoretical basis is documented in

[1] Yang, B., Knyazikhin, Y., Mõttus, M., Rautiainen, M., Stenberg, P., Yan, L., Chen, C., Yan, K., Choi, S., Park, T., & Myneni, R.B. (2017). Estimation of leaf area index and its sunlit

portion from DSCOVR EPIC data: Theoretical basis. *Remote Sensing of Environment*, 198 (https://doi.org/10.1016/j.rse.2017.05.033)

An overview of the DSCOVR EPIC project can be found in

[2] Marshak, A., Herman, J., Szabo, A., Blank, K., Cede, A., Carn, S., Geogdzhayev, I., Huang, D., Huang, L.-K., Knyazikhin, Y., Kowalewski, M., Krotkov, N., Lyapustin, A., McPeters, R., Torres, O., & Yang, Y. Earth Observations from DSCOVR/EPIC Instrument. *Bulletin of the American Meteorological Society*, doi:/10.1175/BAMS-D-17-0223.1 (https://journals.ametsoc.org/view/journals/bams/99/9/bams-d-17-0223.1.xml).

The Directional Area Scattering Factor (DASF) is a new structural parameter that estimates the canopy BRF if the leaves do not absorb radiation. Its definition and analysis of its value for remote sensing of leaf biochemistry are discussed in

[3] Knyazikhin, Y., Schull, M.A., Stenberg, P., Mõttus, M., Rautiainen, M., Yang, Y., Marshak, A., Latorre Carmona, P., Kaufmann, R.K., Lewis, P., Disney, M.I., Vanderbilt, V., Davis, A.B., Baret, F., Jacquemoud, S., Lyapustin, A., & Myneni, R.B. (2013). Hyperspectral remote sensing of foliar nitrogen content. *Proceedings of the National Academy of Sciences*, 110, E185-E192 (https://www.pnas.org/content/110/3/E185)

The Earth Reflector Type Index (ERTI) was developed to discriminate between signals originating from clouds, cloud-free ocean, bare and vegetated land. Its definition and its value for analyses of Top Of Atmosphere (TOA) reflectance are documented in

[4] Song, W.J., Knyazikhin, Y., Wen, G.Y., Marshak, A., Mottus, M., Yan, K., Yang, B., Xu, B.D., Park, T., Chen, C., Zeng, Y.L., Yan, G.J., Mu, X.H., & Myneni, R.B. (2018). Implications of Whole-Disc DSCOVR EPIC Spectral Observations for Estimating Earth's Spectral Reflectivity Based on Low-Earth-Orbiting and Geostationary Observations. *Remote Sensing*, 10 (https://www.mdpi.com/2072-4292/10/10/1594)

Version 2 VESDR product contains ERTI at surface level which is derived from the upstream DSCOVER EPIC L2 MAIAC surface reflectance product.

1.4. Revisions. This document describes version 2 of the VESDR product. It can be downloaded, distributed, and cited. Changes in version 2 VESDR product include (a) projection and product tiling have been changed; (b) 8 new parameters have been added; (c) definition of Quality Assessment variable has been changed.

2. EXPERIMENT OVERVIEW

The Deep Space Climate Observatory (DSCOVR) mission is a multiagency (National Oceanic and Atmospheric Administration [NOAA], U.S. Air Force, and NASA) mission launched from Cape Canaveral, Florida on February 11, 2015 with the primary goal of making unique space weather measurements from the first Sun-Earth Lagrange point (L1). The L1 point is on the direct line between Earth and the Sun located 1.5 million km sunward from Earth. The spacecraft is orbiting this point in a six month Lissajous orbit with a Sun-Earth-View (SEV) angle varying between 4.5° and 11.5°. The primary science objective of the DSCOVR mission is to provide solar wind thermal plasma and magnetic field measurements to enable space weather forecasting by NOAA.

The DSCOVR hosts NASA Earth-Observing Instrument, the Earth Polychromatic Imaging Camera (EPIC). The EPIC provides measurements of the radiation reflected by Earth in ten wavelengths and images of the sunlit side of Earth for science applications.

2.1. EPIC instrument characteristics. The EPIC instrument collects multispectral data of the Earth in ten wavelengths. The spectral band characteristics are summarized in Table 3.

Table 3: EPIC spectral band composition

Wavelength, nm	FWHM,nm	Nominal Product
317.5±0.1	1±0.2	Ozon
325±0.1	2±0.2	Ozon
340±0.3	3±0.6	Ozon, Aerosols, Clouds
388±0.3	3±0.6	Aerosols, Clouds
443±1	3±0.6	Aerosols
551±1	3±0.6	Aerosols, Vegetation
680±0.2	2±0.4	Aerosols, Vegetation, Clouds,O_2 B-Band Reference
687.75±0.2	0.8±0.2	O_2 B-Band Cloud Height
764±0.2	1±0.2	O_2 A-Band Cloud Height, Aerosol Height
779±0.3	2±0.4	O_2 A-Band Reference, Vegetation

2.2. Rationale for the DSCOVR EPIC VESDR product. Fraction vegetation absorbed Photosynthetically Active Radiation (FPAR), Leaf Area Index (LAI), its sunlit counterpart (SLAI), and Normalized Difference Vegetation Index (NDVI are useful for (a) monitoring variability and change in global vegetation due to climate and anthropogenic influences, (b) modeling climate, carbon and water cycles, and (c) improving forecasting of near surface weather. The Directional Area Scattering Factor provides information critical to accounting for structural contributions to measurements of leaf biochemistry from remote sensing. The canopy scattering coefficient is strongly correlated with leaf optical properties, which in turn convey information about leaf biochemical constituents. Whereas LAI is a standard product of many satellite missions, global diurnal courses of FPAR, NDVI, SLAI and DASF are new satellite derived products.

3. SUN-SENSOR GEOMETRY

The sun-sensor geometry is expressed in a right-handed coordinate system in which the Z-axis (shown as "+Z" in Fig. 1) is aligned with the normal to the surface reference ellipsoid (defined by the World Geodetic System 1984, WGS84), and points toward the center of the Earth. The X-axis is aligned with a great circle and points toward the north pole. The Y-axis is orthogonal to both of them.

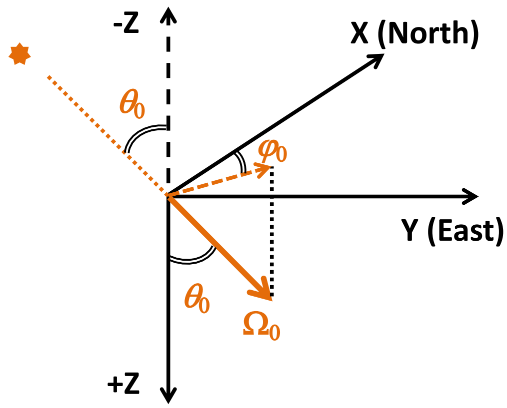

Figure 1. Right-handed coordinate system in which the Z-axis (shown as "+Z") points toward the center of the Earth. The X-axis and Y-axis point toward the North and East, respectively. The direction (unit vector) Ω_0 has an azimuthal angle, φ_0, measured clockwise from the local north vector (X) to the projection of Ω_0 onto the XY plane, and a polar angle, θ_0, with respect to the +Z axis.

The Sun and sensor positions are represented by unit vectors Ω_0 and Ω directed downward from the Sun to target (i.e., point at the Earth surface) and from sensor to target, respectively. Their polar (θ_0 and θ) and azimuthal (φ_0 and φ) angles are given in the right-handed coordinate system (Fig. 1). Their ranges are between 0 and 90° (polar angles) and between 0 and 360° (azimuthal angles). This coordinate system is inherited from the upstream DSCOVER EPIC L2 MAIAC surface reflectance product, which is input to the VESDR retrieval algorithm. In this coordinate system the Solar Zenith Angle (SZA, the angle between the target-to-Sun direction, $-\Omega_0$, and the –Z axis) coincides with the polar angle of Ω_0, i.e., $\theta_0 =$ SZA.

The Earth-observing geometry of the EPIC instrument is characterized by a nearly constant phase angle (the angle between directions to the Sun and to the sensor) between 4.5° and 11.5°. The phase angle, γ, can be calculated as

$$\gamma = \mathrm{acos}(\cos\theta\cos\theta_0 + \sin\theta\sin\theta_0\cos(\varphi - \varphi_0)). \qquad (1)$$

4. PROJECTION AND PRODUCT TILING

The globe is divided into 4 zones according to longitude (Fig. 2). The zones are extracted from 4 rotated sinusoidal projections with central meridians at 20, 110, -160 and -70 degree. Definitions of zones are given in Table 4.

Table 4. Zones of V2 VESDR biome map

Zone ID	Central meridian of SIN projection	Meridian of left bound	Meridian of right bound
Zone 0	20	-25	65
Zone 1	110	65	155
Zone 2	-160	155	-115
Zone 3	-70	-115	-25

Each zone is divided into northern (block number 0) and southern (block number 1) tiles. Thus, the VESDR parameters are projected on 8 regional tiles. Each tile is identified by its zonal Z (from 0 to 3) and block B (0 or 1) coordinates, e.g., tile30.

Dimension of tile: 1000x1000; Spatial resolution: 10018.7542 meter.

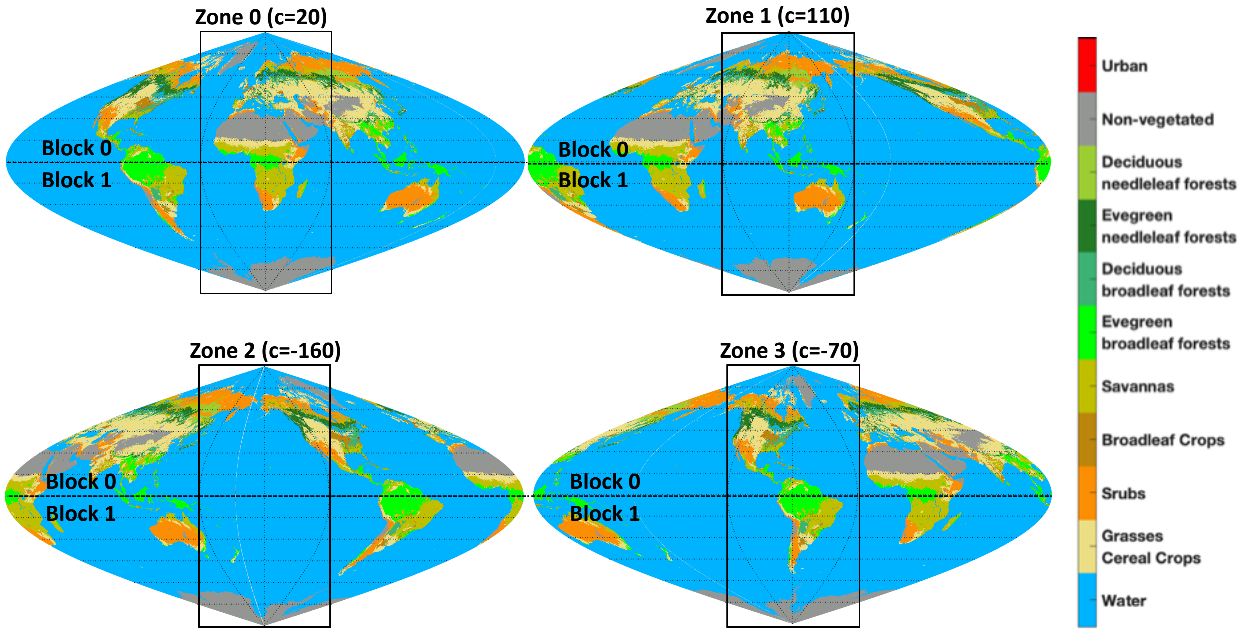

Figure 2. Rotated sinusoidal projections with central meridians at 20, 110, -160 and -70 degree. Four zones are extracted from these maps. Each zone is divided into Northern (block number 0) and Southern (block number 1) tiles. Each tile is identified by its zonal Z (from 0 to 3) and block B (0 or 1) coordinates, e.g., tile30.

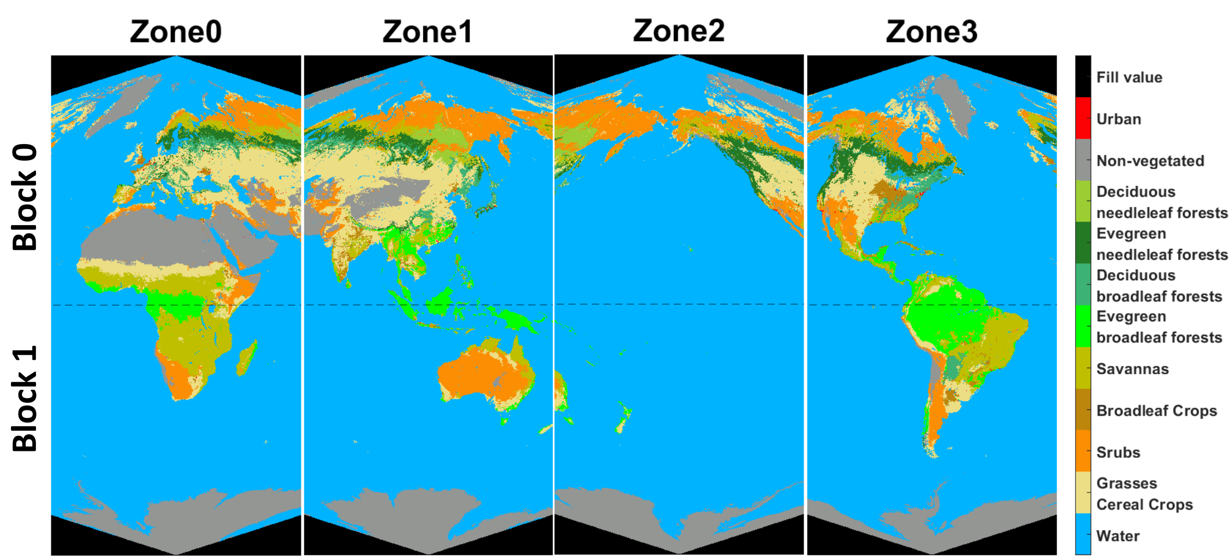

Figure 3. DSCOVR EPIC land cover types projected on 8 regional tiles. Each tile is identified by its zonal Z (from 0 to 3) and block (0 or 1) coordinates, e.g., tile30. The VESDR algorithm uses this ancillary science data product. File name: MCDLCHKM.V2010_01.REGIONAL.10018m.BlocksOP.h5.

5. LEVEL 2 VESDR PRODUCT

5.1. VESDR product file name

The file name containing VESDR parameters is

DSCOVR_EPIC_L2_VESDR_V1_YYYYMMDDHHMMSS_V2.h5

Here V1 and V2 are versions of the VESDR product and L2B TOA reflectance data, respectively. Current versions are V1=02 and V2=03. YYYYMMDDHHMMSS signifies date and GMT time of EPIC image acquisition. For example, file DSCOVR_EPIC_L2_VESDR_02_20160823141930_03.h5 contains VESDR parameters for an EPIC image acquired on August 23, 2016 (20160823) at 14h10m30s GMT (141930).

5.2. HDF file structure

The VESDR product is distributed as standard Hierarchical Data Format 5 (HDF5) file. The data are compressed using the lossless gzip option provided by the HDF5 FORTRAN Application Programming Interface (API). Compression level is 4. File size depends on number of available tiles and varies between 28 and 37 megabytes (MB).

In the HDF5 file, data are grouped by tiles. Each group contains geophysical parameters, associated quality assessment variables (QA_VESDR and Dlai) and sun-sensor geometry. The root directory contains a set of attributes that describes the content of the HDF5 file. Figure 4 illustrates a snapshot of the layout of a L2 VESDR product.

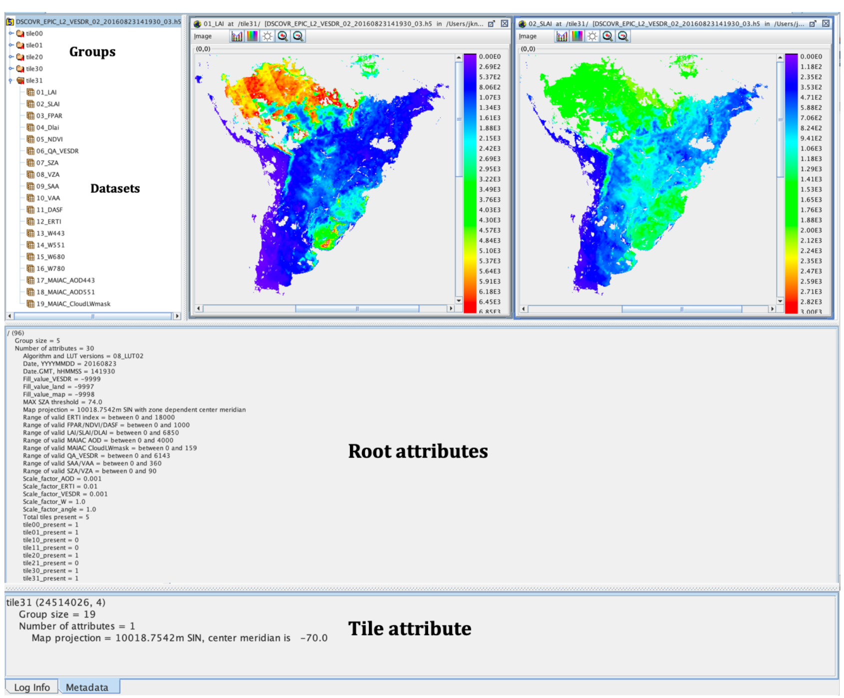

Figure 4. Structure of the L2 VESDR product. Data are grouped by tiles. Each group contains geophysical parameters, associated quality assessment variables and sun-sensor geometry. The root attributes provide general information about the VESDR product. The tile attribute 'projection' provides value for the center meridian.

5.3. Root and tile attributes

Root attributes include algorithm and LUT versions, date and time of the EPIC image acquisition, fill values, map projection, parameter's valid ranges, scale factors and list of tiles present in the file. Details are summarized in Table 5. The tile attribute 'Map projection' provides value of the center meridian.

Table 5: L2 VESDR root attributes

Attribute name	Value, range	Type	Description
Algorithm and LUT versions	08_LUT02	string	Algorithm version 08, LUT version 02
Date, YYYYMMDD	20160613-current date	32 bit integer	Date of EPIC image acquisition
Date.GMT, hHMMSS	00000 - 235959	32 bit integer	GMT of EPIC image acquisition
Fill_value_VESDR	-9999	16 bit integer	VESDR parameter was not generated.
Fill_value_land	-9997	16 bit integer	Non-vegetated pixel
Fill_value_map	-9998	16 bit integer	Out of map boundary ("black area" in Fig. 3)
Max SZA threshold	74.0	32 bit floating point	VESDR algorithm does not process pixel if the SZA exceeds Max_SZA_threshold
Map projection	10018.7542m SIN with zone dependent center meridian	string	Center meridian for each tile is found in tile attribute.
Range of valid ERTI index	between 0 and 18000	string	Range of valid DN values of ERTI index; must be multiplied by Scale_factor_ERTI to get physical value
Range of valid FPAR/NDVI/DASF	between 0 and 1000	string	Range of valid DN values of FPAR, NDVI and DASF; must be multiplied by Scale_factor_VESDR to get physical value
Range of valid LAI/SLAI/DLAI	between 0 and 6850	string	Range of valid DN values of LAI, SLAI and Dlai; must be multiplied by Scale_factor_VESDR to get physical value
Range of valid MAIAC AOD	between 0 and 4000	string	Range of valid DN values of AOD; must be multiplied by Scale_factor_AOD to get physical value
Range of valid MAIAC CloudLWmask	between 0 and 159	string	Range of valid values of Cloud and Land-Water masks
Range of valid QA_VESDR	between 0 and 6143	string	Range of valid values of Cloud and Land-Water masks
Range of valid SAA/VAA	between 0 and 360	string	Range of valid values of solar and view azimuthal angles
Range of valid SZA/VZA	between 0 and 980	string	Range of valid values of solar and view zenith angles
Scale_factor_AOD	0.001	32 bit floating point	AOD DN values at 443nm and 551 nm should be multiplied by the scale factor to convert their DN values to physical value
Scale_factor_ERTI	0.01	32 bit floating point	ERTI DN value should be multiplied by the scale factor to convert their DN values to physical value
Scale_factor_VESDR	0.001	32 bit floating point	DN values of LAI, SLAI, FPAR, NDVI, DASF and Dlai, should be multiplied by the scale factor to convert their DN value to physical value
Scale_factor_W	1.0	32 bit floating point	The VESDR file contains physical values of the scattering coefficient
Scale_factor_angle	1.0	32 bit floating point	The VESDR file contains physical values of sun-sensor geometry parameters

Total tiles present	1-8	8 bit integer	Total number of tiles present in the VESDR file
tile00_present	0,1	8 bit integer	Indicates if tile00 present (value=1) or not (value 0) in the VESDR file
tile01_present	0,1	8 bit integer	Indicates if tile01 present (value=1) or not (value 0) in the VESDR file
tile10_present	0,1	8 bit integer	Indicates if tile10 present (value=1) or not (value 0) in the VESDR file
tile11_present	0,1	8 bit integer	Indicates if tile11 present (value=1) or not (value 0) in the VESDR file
tile20_present	0,1	8 bit integer	Indicates if tile20 present (value=1) or not (value 0) in the VESDR file
tile21_present	0,1	8 bit integer	Indicates if tile21 present (value=1) or not (value 0) in the VESDR file
tile30_present	0,1	8 bit integer	Indicates if tile30 present (value=1) or not (value 0) in the VESDR file
tile31_present	0,1	8 bit integer	Indicates if tile31 present (value=1) or not (value 0) in the VESDR file

5.4. Datasets

5.4.1. Parameters archived in VESDR file. Each tile contains geophysical parameters, associated quality assessment variables (QA_VESDR and Dlai) and sun-sensor geometry. Description of the datasets is given in Table 6.

Table 6: L2 VESDR datasets

Name of dataset	Valid range	Data type	Description
01_LAI	0-6850	16 bit integer	Leaf Area Index
02_SLAI	0-6850	16 bit integer	Sunlit Leaf Area Index
03_FPAR	0-1000	16 bit integer	fraction of photosynthetically active radiation (400 – 700nm) absorbed by vegetation
04_Dlai	0-6850	16 bit integer	Precision of Leaf Area Index
05_NDVI	0-1000	16 bit integer	Normalized Difference Vegetation Index
06_QA_VESDR	0-6143	16 bit unsigned integer	Quality Assessment variable. See section 5.5
07_SZA	0-90	32 bit floating point	Polar angle (in DEG) of the Sun-to-target direction as defined in Sect. 3
08_VZA	0-90	32 bit floating point	Polar angle (in DEG) of the sensor-to-target direction as defined in Sect. 3
09_SAA	0-360	32 bit floating point	Azimuthal angle (in DEG) of the Sun-to-target direction as defined in Sect. 3
10_VAA	0-360	32 bit floating point	Azimuthal angle (in DEG) of the sensor-to-target direction as defined in Sect. 3
11_DASF	0-1000	16 bit integer	Estimate of Canopy Bidirectional Reflectance Factor as if the foliage does not absorb radiation
12_ERTI	0-18000	16 bit integer	Estimate of Earth Reflector Type Index (ERTI)
13_W443	0-1	32 bit floating point	Scattering coefficient at 443 nm
14_W551	0-1	32 bit floating point	Scattering coefficient at 551 nm
15_W680	0-1	32 bit floating point	Scattering coefficient at 680 nm
16_W780	0-1	32 bit floating point	Scattering coefficient at 780 nm
17_MAIAC_AOD443	0-4000	16 bit integer	Aerosol Optical Depth at 443 nm
18_MAIAC_AOD443	0-4000	16 bit integer	Aerosol Optical Depth at 551 nm
19_MAIAC_CloudLWmask	0-159	8 bit unsigned integer	Cloud and Land-Water masks. See section 5.4.2 for details

5.4.2. Cloud and Land-Water masks. Parameter 19_MAIAC_CloudLWmask contains information about cloud state (bits 0 to 3), land and water mask (bits 4 to 7). It is archived as 8 bit unsigned integer number. These parameters were copied from upstream DSCOVR EPIC L2 MAIAC (version 2) product. Details are given in Table 7.

Table 7: Cloud and Land-Water masks (19_MAIAC_CloudLWmask)

	Binary value	Decimal value	Description
Bits 0-3: MAIAC cloud mask	0000	0	Value 255 or 50 in MAIAC cloud mask
	0001	1	Clear
	0010	2	Definition is not provided
	0011	3	Possibly cloud
	0100	4	Cloud
	0101	5	Definition is not provided
	0110	6	Cloud shadow
	0111	7	Clear, with smoke detected
	1000	8	Clear, with dust detected
	1001	9	Clear, over water, with water sediments detected
	1010	10	clear, water
	1011	11	Definition is not provided
	1100	12	Definition is not provided
	1101	13	Glint
	1110	14	Definition is not provided
	1111	15	Misclassified land/water
Bits 4-7: MAIAC Land-Water mask	0000	0	Not defined
	0001	1	Land
	0010	2	Land
	0011	3	Land
	0100	4	Snow
	0101	5	Ice over water
	0110	6	General water
	0111	7	Deep water
	1000	8	Shallow water
	1001	9	Static sea
	1010	10	Static lake
	1011	11	Not defined
	1100	12	Not defined
	1101	13	Not defined
	1110	14	Not defined
	1111	15	Value 255 in MAIAC Land Water Mask

5.5. Quality assessment dataset

Version 2 VESDR product is being generated from the upstream version 2 DSCOVER EPIC L2 MAIAC surface reflectance product. Quality of the VESDR parameters depends on quality of the MAIAC data. Uncertainties in the surface BRFs are high if aerosol optical depths at 443 nm and 551nm exceed 0.6 and 0.3, respectively. Therefore we recommend consult the aerosol optical depth and use VESDR parameters retrieved when AOD443<0.6 and AOD551<0.3.

5.5.1. Information content of QA dataset. Quality assessment variable, 06_QA_VESDR, includes quality control information on VESDR algorithm performance (bits 0 to 8) and Status_QA (bits 9 to 12). The latter is copied from the upstream version 2 DSCOVR EPIC L2

MAIAC surface reflectance product. The DSCOVR EPIC MAIAC product is input to the VESDR retrieval technique. 06_QA_VESDR therefore provides information about quality of both input to the VESDR algorithm and the VESDR algorithm output. Figure 5 shows structure of 06_QA_VESDR. Details are given in Table 8.

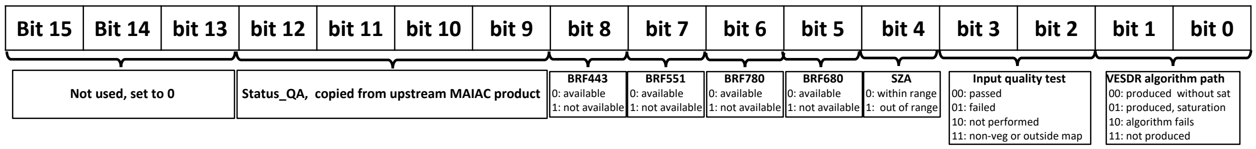

Figure 5 Information content of 16 bit unsigned integer 06_QA_VESDR

Table 8: Values of 16 bit unsigned integer 06_QA_VESDR

QA name	Bits	Binary value	Decimal value	Description
VESDR algorithm path	0-1	00	0	LAI, SLAI, FPAR, DLAI produced by main algorithm without saturation
		01	1	LAI, SLAI, FPAR, DLAI produced by main algorithm under saturation conditions
		10	2	LAI, SLAI, FPAR, DLAI were not found for given surface reflectance
		11	3	LAI, SLAI, FPAR, DLAI were not produced because (a) pixel is not non-vegetated; or (b) red or nir BRF not available or negative. See other bits
Input quality test	2-3	00	0	Input quality test passed
		01	1	Input quality test failed
		10	2	input quality test not perform because NIR or GREEN BRF not available, DASF, ERTI and W were not generated
		11	3	VESDR parameters were not produced because of outside the map
SZA	4	0	0	SZA is between 0^{o} and maxSZAthreshod. See root attributes.
		1	1	LAI, SLAI, FPAR, DLAI were not produced because SZA is out of range or out of map
Availability of BRF	5	0	0	BRF 680 nm available
		1	1	BRF 680 nm not available or negative
	6	0	0	BRF 780 nm available
		1	1	BRF 780 nm not available or negative
	7	0	0	BRF 551 nm available
		1	1	BRF 551 nm not available or negative
	8	0	0	BRF 443 nm available
		1	1	BRF 443 nm not available or negative
Status_QA from upstream MAIAC product	9-12	0000	0	Best quality
		0001	1	Clear water, sediments detected
		0010	2	1 neighbor cloud
		0011	3	>1 neighbor clouds
		0100	4	no retrieval (cloudy, or whatever)
		0101	5	definition is not provided
		0110	6	for H>3.5km, use climatology AOD=0.2
		0111	7	definition is not provided
		0100	8	sun glint
		1001	9	land-water misclassified
		1010	10	no retrievals: CoxMunk too high
		1011	11	"255" in MAIAC file; or Status_QA < 0 or Status_QA >10

5.5.2. Saturation, Retrieval Index and input quality test. In the case of dense canopies, the reflectances saturate and therefore are weakly sensitive to changes in canopy properties. The reliability of parameters retrieved under the condition of saturation is low. Such retrievals are flagged by setting decimal value of the VESDR_algorithm_path to 1 (binary value ='01').

The retrieval index, RI, is the percentage of pixels with valid BRF for which the VESDR algorithm produced a retrieval. The index characterizes the spatial coverage of the geophysical parameters. This important characteristic of the algorithm performance can be calculated as

$$RI = \frac{N(\text{bits01} = '\ 00' \text{or bits01} = '\ 01')}{N(\text{bit5} = '\ 0' \text{ and bit6} = '\ 0')}. \tag{2}$$

Here the numerator represents number of pixels for which the VESDR_algorithm_path is 0 or 1. The denominator is the number of pixels for which values of bit 5 and bit 6 are 0.

For vegetated pixels at weakly absorbing wavelengths, the BRF to leaf albedo ratio is linearly related to BRF, i.e.,

$$\frac{BRF_\lambda}{\omega_\lambda} = pBRF_\lambda + R\ , \tag{3}$$

where the slope, p, and intercept, R, are the recollision probability and escape factor. We use BRF at green and NIR EPIC bands to estimate the slope, p, of a line passing points $\left(\frac{BRF_{green}}{\omega_{green}}, BRF_{green}\right)$ and $\left(\frac{BRF_{NIR}}{\omega_{NIR}}, BRF_{NIR}\right)$ on the $\frac{BRF}{\omega}$ vs BRF plane. Its value is given by

$$p = \frac{\frac{BRF_{NIR}}{\omega_{NIR}} - \frac{BRF_{green}}{\omega_{green}}}{BRF_{NIR} - BRF_{green}}. \tag{4}$$

Here ω_λ represents a reference leaf albedo at NIR and green spectral bands [publ. 3 in Sect 1.3]. Their values at these bands are set to 0.4898 (green) and 0.9789 (NIR) in the VESDR operational algorithm. Our analyses suggest that Eq. (4) takes values between 0 and 1 only for vegetated surfaces [publ. 4 in Sect. 1.3]. For BRF at green and NIR spectral bands over non-vegetated land, water or cloud-contaminated pixels, Eq.(4) generates values outside of the 0 to 1 range. This property underlies the input quality test: bits 2-3 are set to '00' if p is between 0 and 1, and to '01', otherwise. The VESDR algorithm processes pixels irrespective of the test result. The input_quality_test QA is just a warning that VESDR parameters were retrieved using input BRF of suspicious quality.

The Earth Reflector Type Index (ERTI) is the recollision probability p transformed to the interval [0°,180°] as $\text{atan}(p)$ if $\text{atan}\, p \geq 0$ and $\text{atan}(p) + 180°$ otherwise. Since p is the slope of the $BRF_\lambda/\omega_\lambda$ vs BRF_λ line, the ERTI is just an angle between this line and BRF_λ axis. Middle panel of Figure 6 shows ERTI values derived from an EPIC Top Of Atmosphere (TOA) reflectance data (left panel). One can see that TOA ERTI corresponding to cloud free land, vegetation, ocean and cloudy pixels tend to occupy different spaces within the 0° to 180° interval. The VESDR product provides ERTI values at surface level generated from atmospherically corrected BRFs at NIR and green spectral bands (right panel).

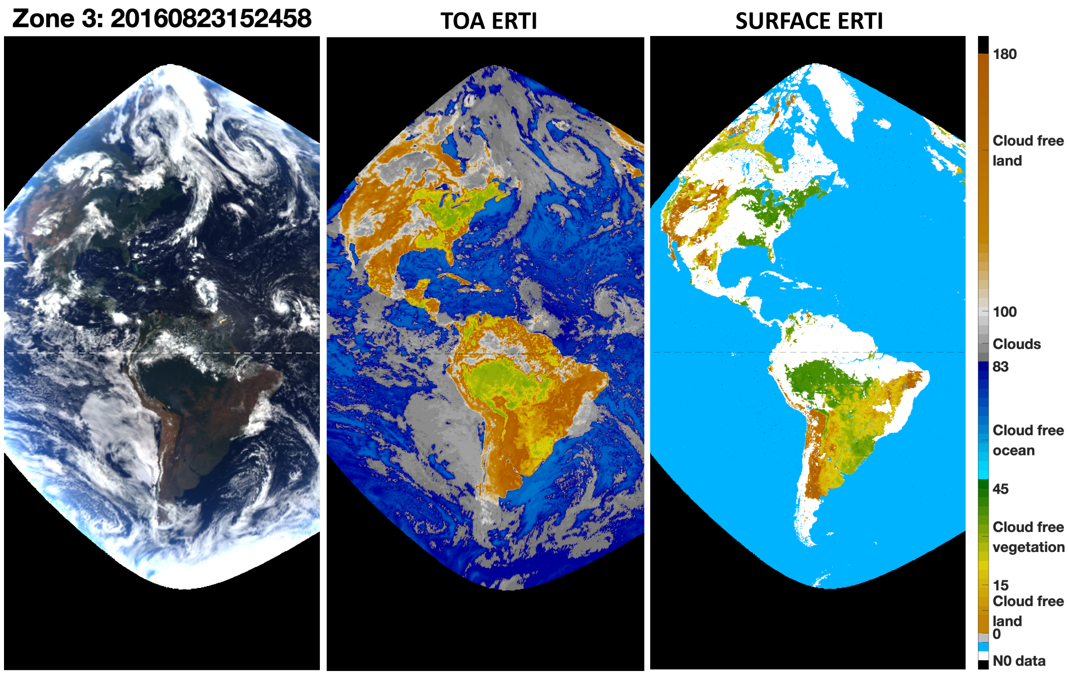

Figure 6. Left panel shows an EPIC RGB image taken on August 23, 2016 at 15:24:58GMT projected on Zone 3 regional grid. Eq. (4) was applied to each image pixel. The slope p was converted to ERTI as $\text{atan}(p)$ if $\text{atan}\, p \geq 0$ and $\text{atan}(p) + 180°$ otherwise. Middle panel shows distribution of ERTI values over the EPIC image. Its values corresponding to cloud free land, vegetation, ocean and cloudy pixels tend to occupy different spaces within the 0° to 180° interval. Right panel shows VESDR ERTI at surface level derived from atmospherically corrected surface BRF.

6. ANCILLARY SCIENCE DATA PRODUCTS

A land cover map is an important ancillary data layer used by the VESDR retrieval algorithm. The global classification of canopy structural types utilized in the Collection 6 MODIS LAI/FPAR algorithm is adopted. Global vegetation is stratified into eight canopy architectural types, or biomes. The eight biomes are Grasses and Cereal Crops (B1), Shrubs (B2), Broadleaf Crops (B3), Savannas (B4), Evergreen Broadleaf Forests (B5), Deciduous Broadleaf Forests (B6), Evergreen Needle Leaf Forests (B7) and Deciduous Needle Leaf Forests (B8).

The VESDR ancillary science data products include *10 km Land Cover Type* and *Distribution of Land Cover Types within 10 km EPIC pixel*. These products were derived from the MODIS 8-biome SIN 500 m resolution land cover type 3 product (MCDLCHKM), which was generated from 2008, 2009 and 2010 MODIS land cover products (MCD12Q1, v051).

6.1. DSCOVR EPIC land cover type

The MODIS Land Cover Product is projected on 500 m sinusoidal (SIN) grid. A 10 km EPIC SIN grid pixel therefore contains about 400 MODIS pixels with known land cover types. The EPIC land cover type is assigned based on the dominant land cover fraction. If there are several land cover types with equal frequency, biome type with highest biome number BN is taken as the EPIC land cover type. For example, if B5 (Deciduous Broadleaf Forests), B4 (Shrubs) and B1 (Grasses and Cereal Crops) occupy 40%, 40% and 20% of the pixel area, then B5 is assigned to the EPIC land cover type. The most frequent land cover type numbers are also stored in the DSCOVR EPIC land cover file. Figure 3 shows eight tiles of 10018.7542m SIN DSCOVR EPIC land cover type.

<u>DSCOVR EPIC land cover product file name:</u> MCDLCHKM.V2010_01.REGIONALbio.10018m.BlocksOP.h5

6.1.1. HDF file structure. In the HDF5 file, data are grouped by tiles. Each root group contains two data sets, "Land_Cover_Type_3" and "Multi_Land_Cover_Types_presented," as well as a tile group "Geolocation" with two datasets, "Latitude" and "Longitude." The root group directory contains a set of attributes that describes the content of the HDF5 file. Each tile has attribute 'projection', which provides value of the center meridian. Figure 7 shows a snapshot of the HDFView layout of MCDLCHKM.V2010_01.REGIONALbio.10018m.BlocksOP.h5 file.

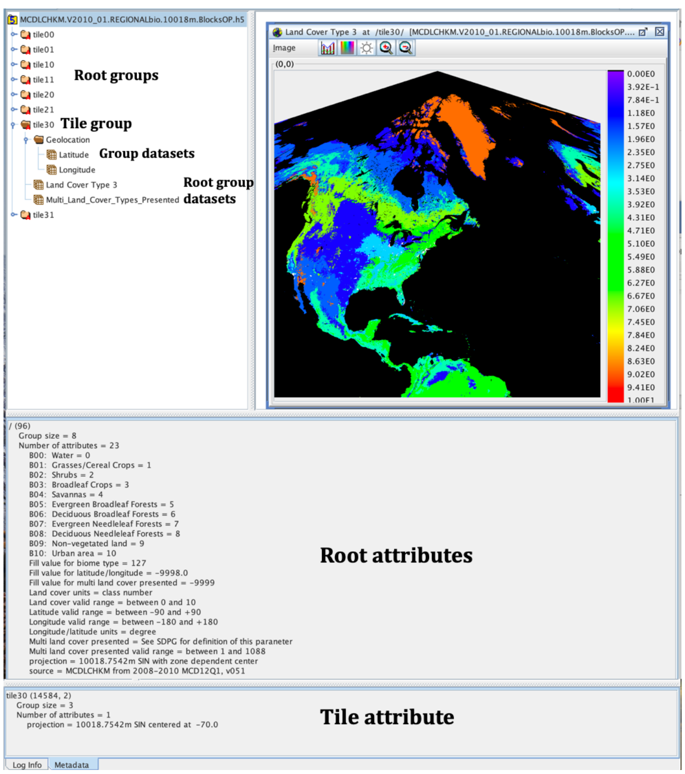

Figure 7. Structure of the ancillary 10018.7542m EPIC Land Cover Type Product

6.1.2. Root and tile attributes. Root attributes include land cover type IDs and associated values, fill values, parameter valid ranges, units, map projection info and short description of the "Multi_Land_Cover_Types_presented" dataset. The tile attribute 'projection' provides value of the center meridian. Details are summarized in Table 9.

Table 9: Land Cover type root attributes

Attribute name	Value	Type	Description
00_Water	0	8 bit integer	Pixel is classified as water
01_Grasses/Cereal Crops	1	8 bit integer	Pixel is classified as B1: Grasses and Cereal Crops
02_Shrubs	2	8 bit integer	Pixel is classified as B2: Shrubs
03_Broadleaf Crops	3	8 bit integer	Pixel is classified as B3: Broadleaf Crops
04_Savannas	4	8 bit integer	Pixel is classified as B4: Savannas
O5_Evergreen Broadleaf Forests	5	8 bit integer	Pixel is classified as B5: Evergreen Broadleaf Forests
06_Deciduous Broadleaf Forests	6	8 bit integer	Pixel is classified as B6 : Deciduous Broadleaf Forests
07_Evergreen Needle Leaf Forests	7	8 bit integer	Pixel is classified as B7: Evergreen Needle Leaf Forests
08_Deciduous Needle Leaf Forests	8	8 bit integer	Pixel is classified as B8: Deciduous Needle Leaf Forests
09_Non-vegetated land	9	8 bit integer	Pixel is classified as non-vegetated land
10_Urban area	10	8 bit integer	Pixel is classified as urban area
Fill value for biome type	127	8 bit integer	Out of map pixel ("Black area" in Fig. 3)
Fill value for latitude/longitude	-9998.0	32 bit floating point	Out of map pixel
Fill value for multi land cover presented	-9999	16 bit integer	Out of map pixel
Land cover units	class number	string	Land cover type ID
Land cover valid range	between 0 and 10	string	Valid range in the "Land_Cover_Type_3" dataset
Latitude valid range	between -90 and +90	string	Valid range in the "Latitude" dataset
Longitude valid range	between -180 and +180	string	Valid range in the "Latitude" dataset
Longitude/ Latitude units	degree	string	Units of latitude and longitude
Multi land cover types presented	See SDPG for definition of this parameter	string	Description of the "Multi_land_cover_types presented" dataset. See Sect. 6.1.4.
Multi land cover types presented valid range	between 1 and 1088	string	Valid range in "Multi_land_cover_types presented" dataset
Projection	10018.7541m SIN with zone dependent center (type: string)		Center meridian for each tile is found in tile attribute.
Source	MCDLCHKM from 2008-2010 MCD12Q1, V051 (type: string)		Data sets from which land cover types were derived

6.1.3. Datasets. Each root group contains two data sets, "Land_Cover_Type_3" and "Multi_Land_Cover_Types_presented," as well as tile group "Geolocation" with two datasets, "Latitude" and "Longitude" (Fig. 7). Each tile group has attribute "projection". Its value shows center meridian. The values are: "projection=10018.7541m SIN centered at 20" in tiles 00 and 01; "projection=10018.7541m SIN centered at 110" in tiles 10 and 11; "projection=10018.7541m SIN centered at -160" in tiles 20 and 21 and "projection=10018.7541m SIN centered at -70" in tiles 30 and 31. Description of the datasets are given in Table 10.

Table 10: Land Cover type datasets

Name of dataset	Valid range	Data type	Description
Land_Cover_Type_3	between 0 and 11	8 bit integer	Land cover type
Multi land cover types presented	between 1 and 1088	8 bit integer	Information about multiple land cover types with equal frequency. See Sect. 6.1.4
Latitude	between -90 and +90	32 bit floating point	Latitude
Longitude	between -180 and +180	32 bit floating point	Longitude

6.1.4. Multi land cover types presented. This dataset provides information about dominant land cover types within 10 km EPIC pixels. This information is stored as 16-bit integer number. Bits 0 to 10 represent land cover type (Fig. 8), with bit value 1 indicating dominant land cover type. For example, if B0 (Water), B2 (Shrubs), B4 (Savannas) and B5 (Evergreen Broadleaf Forests) occupy 10%, 30%, 30% and 30% of the pixel area, then Multi_Land_Cover Types_Presented is '110100'=52. Bit 11 with bit value 1 indicates fill value, i.e., land cover type was not identified. In this case Multi_Land_Cover Types_Presented is '100000000000'=4095.

Bit 15	Bit 14	bit 13	bit 12	bit 11	bit 10	bit 9	bit 8	bit 7	bit 6	bit 5	bit 4	bit 3	bit 2	bit 1	bit 0
0	0	0	0	Fill value	urban	Non-vegetated	B8	B7	B6	B5	B4	B3	B2	B1	water

Figure 8. Structure of Multi_Land_Cover Types_Presented

6.2. Distribution of land cover types.

Distribution of land cover types within 10 km EPIC pixel is defined as

$$LC_i = 100\% \frac{N_i}{N}. \quad (5)$$

Here i (i=0,1,2,…,10) represents land cover type ID, N_i is the number of the ith land cover type and N is the total number of pixels; $\sum_0^{10} LC_i = 100\%$. Sets of the MODIS 500 m land cover type 3 product within 10 km EPIC pixel was used to derive this distribution.

DSCOVR EPIC land cover product file name is MCDLCHKM.V2010_01.REGIONALdst.10018m.Zones.h5.

6.2.1. HDF file structure. In the HDF5 file, data are grouped by zones. Each zone contains distribution of land cover type within 10 km EPIC pixel. The root group directory contains a set of attributes that describes the content of the HDF5 file. Each zone has attribute 'Map projection', which provides value of the center meridian. Figure 9 illustrates a snapshot of the HDFView layout of MCDLCHKM.V2010_01.REGIONALdst.10018m.Zones.h5 file.

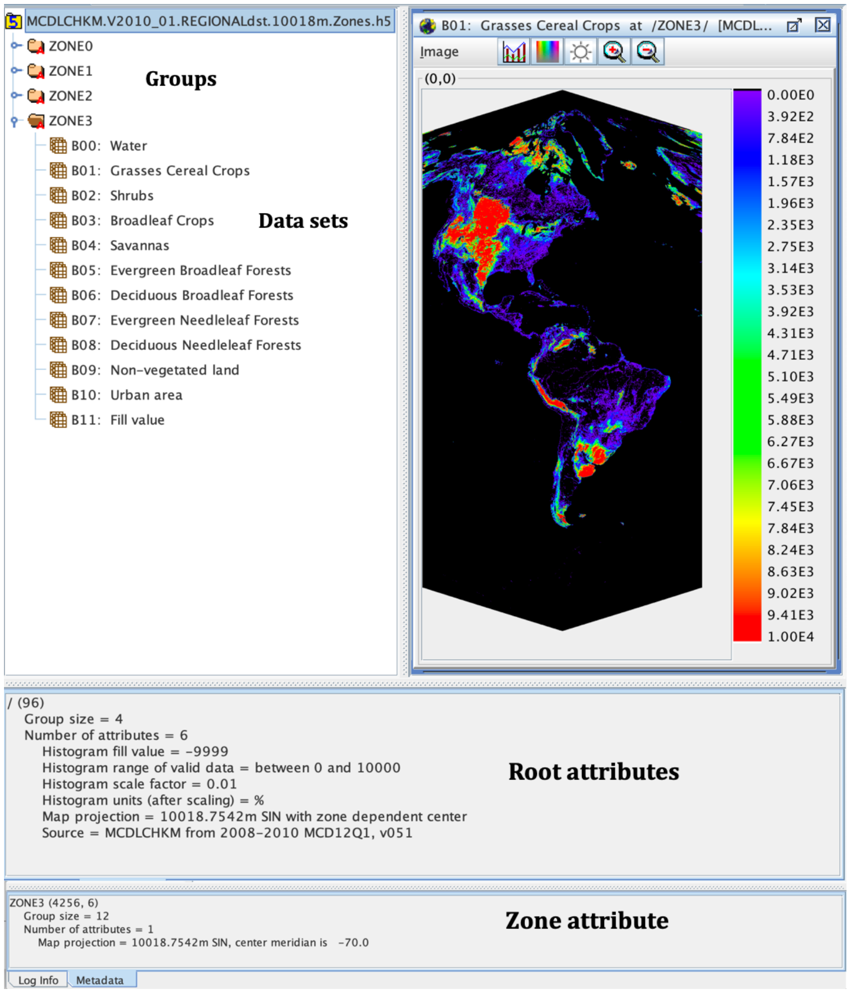

Figure 9. Structure of the ancillary 10018.7542m SIN EPIC Land Cover Type Distribution

6.2.2. Root and Zone attributes. Root attributes include fill value, scale factor, units and distribution valid range. Details are summarized in Table 10.

Table 10: Land cover type distribution root attributes

Attribute name	Value	Type	Description
Histogram fill value	-9999	16 bit integer	Fill value
Histogram valid range	Between 0 and 10000	string	Valid range histogram DN values
Histogram scale factor	0.01	32 bit floating point	Distribution dataset should be multiplied by the scale factor to convert its DN value to physical value
Histogram units	%	string	Units of the distribution
Map projection	10018.7541m SIN with zone dependent center (type: string)		Projection type. Center meridians are given in ZONE's attributes
Source	MCDLCHKM from 2008-2010 MCD12Q1, V051 (type: string)		Data sets from which land cover types were derived

6.2.3. Datasets. Each group contains distribution of land cover type. Each zone has attribute "Map projection". Its value shows center meridian: "Map projection =10018.7541m SIN centered at 20" in zone 0; "Map projection=10018.7541m SIN centered at 110" in zone 1;

“Map projection=10018.7541m SIN centered at -160” in zone 2 and “Map projection=10018.7541m SIN centered at -70” in zone 3. Data sets are detailed in Table 11.

Table 11: Land cover type distribution datasets

Name of dataset	Value	Type	Description
00_Water	0-10000	16 bit integer	DN value of percentage of water in EPIC pixel
01_Grasses Cereal Crops	0-10000	16 bit integer	DN value of percentage of B1: Grasses and Cereal Crops in EPIC pixel
02_Shrubs	0-10000	16 bit integer	DN value of percentage of B2: Shrubs in EPIC pixel
03_Broadleaf Crops	0-10000	16 bit integer	DN value of percentage of B3: Broadleaf Crops in EPIC pixel
04_Savannas	0-10000	16 bit integer	DN value of percentage of B4: Savannas in EPIC pixel
O5_Evergreen Broadleaf Forests	0-10000	16 bit integer	DN value of percentage of B5: Evergreen Broadleaf Forests in EPIC pixel
06_Deciduous Broadleaf Forests	0-10000	16 bit integer	DN value of percentage of B6 : Deciduous Broadleaf Forests in EPIC pixel
07_Evergreen Needleleaf Forests	0-10000	16 bit integer	DN value of percentage of B7: Evergreen Needle Leaf Forests in EPIC pixel
08_ Deciduous Needleleaf Forests	0-10000	16 bit integer	DN value of percentage of B8: Deciduous Needle Leaf Forests in EPIC pixel
09_Non-vegetated land	0-10000	16 bit integer	DN value of percentage of non-vegetated land in EPIC pixel
10_Urban area	0-10000	16 bit integer	DN value of percentage of urban area in EPIC pixel
11_Fill value	0-10000	16 bit integer	DN value of percentage of fill values in EPIC pixel

7. KNOWN ISSUES

Underestimation of VESDR parameters at high solar zenith angle. We recommend using VESDR parameters for SZA<55^{o}.

8. EXAMPLES

This section provides examples of obtaining new information on canopy structure from the VESDR product.

LAI and SLAI satisfy the following equation (publ. [1] in sect.1.3)

$$SF = \frac{1 - t_0}{|\ln t_0|}. \qquad \text{(E1)}$$

Here SF=SLAI/LAI is the Sunlit Fraction (SF) of leaf area and t_0 represents direct transmittance. Given t_0, one can estimate canopy interceptance and Fractional Vegetation Cover (FVC) as $i_0 = 1 - t_0$ and $FVC = 1 - t_0^{\mu}$, respectively. Here $\mu = \cos SZA$. The Clumping Index (CI) is a measure of foliage aggregation relative to a random distribution of leaves in space. This variable is obtained by fitting Beer’s exponential transmission law to measured canopy direct transmittance, i.e., $t_0(\mu) \sim \exp(-G(\mu) \cdot CI \cdot LAI)/\mu)$ where G is the geometry factor.

Example 1: Find canopy direct transmittance and interceptance given LAI and SLAI.

The canopy direct transmittance can be found by solving Eq. (E1) for t_0. Figure E1 shows SFLA (left hand side of Eq. E1), canopy direct transmittance (solution of Eq. E1) and Land Cover Type in South America.

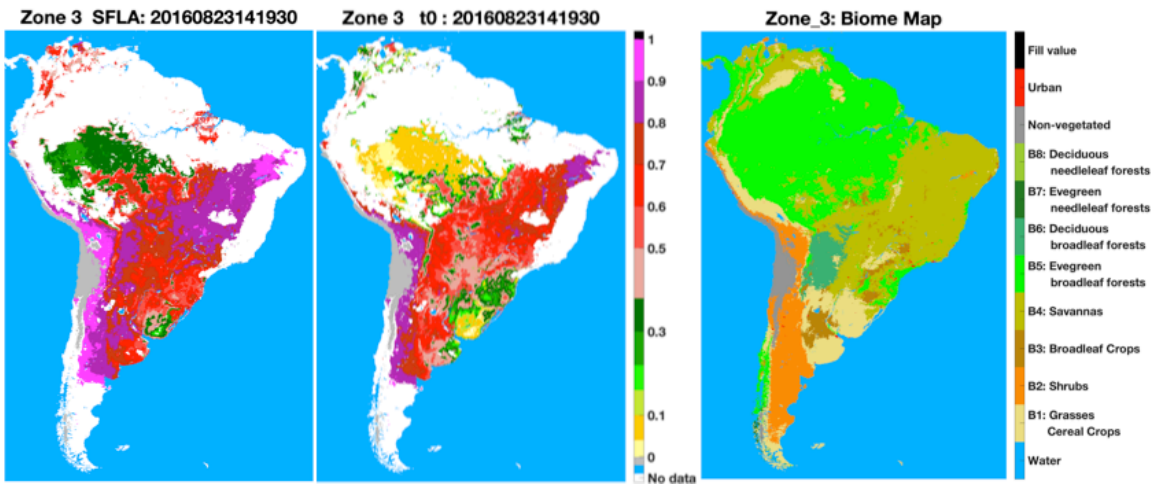

Figure E1. Sunlit Fraction of Leaf Area (left), canopy direct transmittance (center) and biome map (right)

Example 2: Find fractional vegetation cover and clumping index given canopy direct transmittance

Left panel shows fractional vegetation cover calculated as $FVC = 1 - t_0^{\mu}$. Middle plot shows optical path through the canopy calculated as $\tau = -\ln t_0$. Assuming spherical leaf orientation, this variable can be estimated as $\tau = 0.5 \cdot CI \cdot LAI/\mu$. Right panel shows clumping index calculated as $CI = 2\tau\mu/LAI$.

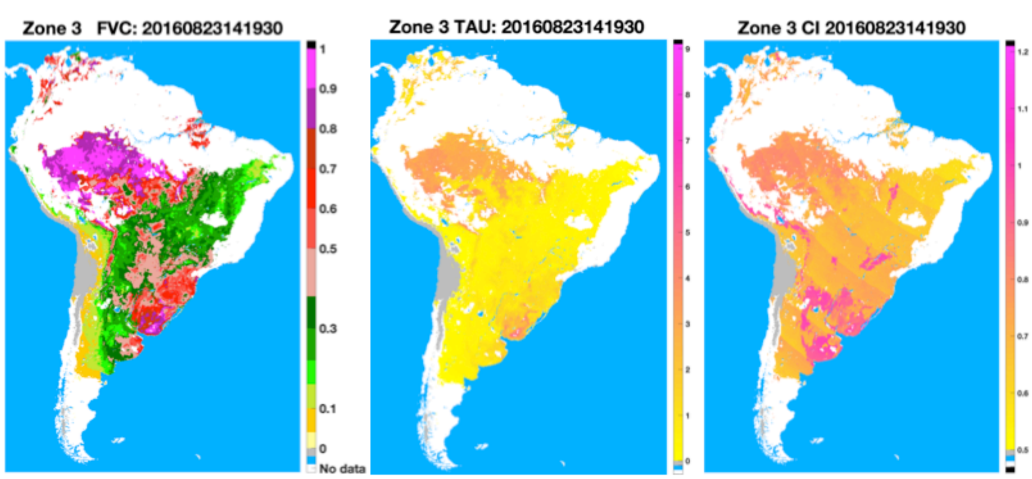

Figure E2. fractional vegetation cover (left), optical path through the canopy (center) and clumping index (right)

Example 3: Find spectral surface BRF_λ given DASF and spectral Scattering Coefficient W_λ.

The surface spectral Scattering Coefficient, W_λ, is calculated as $W_\lambda = BRF_\lambda/DASF$, where BRF_λ, $\lambda = 443{,}551{,}680, 780$ nm, is the atmospherically corrected surface reflectance (parameter BRF_λ in the upstream DSCOVER EPIC L2 MAIAC surface reflectance product). The spectral surface BRF therefore can be calculated as $BRF_\lambda = W_\lambda \cdot DASF$.