SMAP L4 Global Daily 9 km Carbon Net Ecosystem Exchange, Version 4

USER GUIDE

How to Cite These Data

As a condition of using these data, you must include a citation:

Kimball, J. S., L. A. Jones, A. Endsley, T. Kundig, and R. Reichle. 2018. *SMAP L4 Global Daily 9 km EASE-Grid Carbon Net Ecosystem Exchange, Version 4*. [Indicate subset used]. Boulder, Colorado USA. NASA National Snow and Ice Data Center Distributed Active Archive Center. https://doi.org/10.5067/9831N0JGVAF6. [Date Accessed].

FOR QUESTIONS ABOUT THESE DATA, CONTACT NSIDC@NSIDC.ORG

FOR CURRENT INFORMATION, VISIT https://nsidc.org/data/SPL4CMDL

National Snow and Ice Data Center

TABLE OF CONTENTS

1 DATA DESCRIPTION

1.1 Parameters

This SMAP data product contains daily estimates of global ecosystem productivity, including net ecosystem exchange (NEE), gross primary production (GPP), heterotrophic respiration (Rh), and soil organic carbon (SOC), along with quality control metrics. The NEE of CO2 with the atmosphere is a fundamental measure of the balance between carbon uptake by vegetation GPP, and carbon losses through autotrophic respiration (Ra) and heterotrophic respiration (Rh). The sum of Ra and Rh defines the total ecosystem respiration rate (Rtot), which encompasses most of the annual terrestrial CO2 efflux to the atmosphere. All parameters are expressed in units of g C m^{-2} day^{-1}. The CO2 flux state variable outputs are provided in SPL4CMDL files as eight vegetated land-cover classes called Plant Function Types (PFTs). For example, the CO2 flux state variable outputs are provided in *NEE/nee_pft{1..8}_mean*, *GPP/gpp_pft{1..8}_mean*, and *RH/gpp_pft{1..8}_mean*. The soil carbon pool state variable output are provided in *SOC/soc_pft{1..8}_mean*. Refer to Table 1 for descriptions of the eight PFTs.

Table 1. Plant Function Type (PFT) Classifier Summary

PFT Class Label	PFT Code	PFT Description	PFT Class used in SPL4CMDL
Water	0	For all ocean and perennial inland water bodies	No
Evergreen needleleaf	1	Evergreen needle-leaf trees (mostly conifers)	Yes
Evergreen broadleaf	2	Evergreen broadleaf trees	Yes
Deciduous needleleaf	3	Deciduous needle-leaf trees	Yes
Deciduous broadleaf	4	Deciduous broad-leaf trees	Yes
Shrub	5	Shrub (woody perennial)	Yes
Grass	6	Grasses (native Graminoids)	Yes
Cereal crop	7	Cereal cropland (domesticated agricultural crops such as wheat, oats, barley, rye)	Yes
Broadleaf crop	8	Broadleaf crop (domesticated agricultural)	Yes
Urban and Built-up	9	Urban and built-up (cities, towns, highways, etc)	No

PFT Class Label	PFT Code	PFT Description	PFT Class used in SPL4CMDL
Snow and ice	10	Snow and ice (may or may not be perennial)	No
Barren (rock) or sparsely vegetated	11	Barren, rock, or very sparsely vegetated land	No
Unclassified	254	Areas otherwise not classified as per above	No
Totals for each vegetated land class (i.e. count of vegetated 1 km grid cells contained within each 9 km grid cell) are provided in each SPL4CMDL file (QA/qa_count_pft{1..8}). Non-vegetated grid cells are determined by the union of specified vegetation PFT classes in Table 1 and availability of long-term MODIS fPAR (MOD15A2) for production of the fPAR climatology (refer to the Baseline Algorithm). Vegetated PFT grid cells lacking sufficient fPAR retrievals to produce the fPAR climatology and non-vegetated PFT grid cells with otherwise valid fPAR climatology are excluded from SPL4CMDL simulations and QA counts. QA counts are time-static and are therefore identical across files because the PFT classification does not change over the course of data generation within each SPL4CMDL version. Users may use the QA count information to compute total non-vegetated 1 km grid cell coverage, compute percent coverage for each PFT, and account for non-vegetated regions when computing areal averages from SPL4CMDL state variables. For example, when computing the total GPP within a 9 km grid cell, a user would multiply the mean GPP (i.e. /GPP/gpp_mean in g C m^{-2} d^{-1}) by the vegetated PFT total QA count (i.e. /QA/qa_count).			

Refer to the Product Specification Document for details on all parameters.

1.2 Format

Data are in HDF5 format. For software and more information, including an HDF5 tutorial, visit the HDF Group's HDF5 Web site.

1.3 File Contents

As shown in Figure 1, each HDF5 file is organized into the following main groups, which contain additional groups and/or data sets:

Figure 1. Subset of File Contents. For a complete list of file contents for the SMAP Level-4 carbon product, refer to the Product Specification Document.

1.4 Data Fields

Each file contains the main data groups summarized in this section. For a complete list and description of all data fields within these groups, refer to the Product Specification Document.

All global data arrays have dimensions of 1624 rows and 3856 columns (6,262,144 pixels per layer). **Note:** The EASE-Grid 2.0 global 1 km reference grid is defined as 14616 lines by 34704 samples (507,233,664 pixels per layer).

1.4.1 EC

Environmental Constraints Data

1.4.2 GEO

Geolocation data, including latitude/longitude coordinate variables in decimal degree units that enable convenient geo-referenced viewing and analysis.

1.4.3 GPP

Gross Primary Production Data

1.4.4 NEE

Net Ecosystem CO2 Exchange Data

1.4.5 QA

QA includes quality control flags, quality assessment, and valid grid cell counts.

1.4.6 RH

Heterotrophic Respiration Data

1.4.7 SOC

Soil Organic Carbon Data

1.5 Metadata Fields

Includes all metadata that describe the full content of each file. For a description of all metadata fields for this product, refer to the Product Specification Document.

1.6 File Naming Convention

Files are named according to the following convention, which is described in Table 2:

SMAP_L4_C_MDL_yyyymmddThhmmss_VLMmmm_NNN.[ext]

For example:

SMAP_L4_C_mdl_20151007T000000_Vv2020_001.h5

Table 2. File Naming Conventions

Variable	Description
SMAP	Indicates SMAP mission data
L4_C_MDL	Indicates specific product (L4: Level-4; C: Carbon; MDL: Model)

<table>
<tr><th>Variable</th><th>Description</th></tr>
<tr><td>yyyymmddThhmmss</td><td>Date/time in Universal Coordinated Time (UTC) of the first data element that appears in the product, where:
<table>
<tr><td>yyyymmdd</td><td>4-digit year, 2-digit month, 2-digit day</td></tr>
<tr><td>T</td><td>Time (delineates the date from the time, i.e. yyyymmddThhmmss)</td></tr>
<tr><td>hhmmss</td><td>2-digit hour, 2-digit minute, 2-digit second</td></tr>
</table></td></tr>
<tr><td>VLMmmm</td><td>Science Version ID, where:
<table>
<tr><th>Variable</th><th>Description</th></tr>
<tr><td>V</td><td>Version</td></tr>
<tr><td>L</td><td>Launch Indicator (V: Validated Data)</td></tr>
<tr><td>M</td><td>1-Digit Major Version Number</td></tr>
<tr><td>mmm</td><td>3-Digit Minor Version Number</td></tr>
</table>
Example: Vv2020 indicates a Validated product with a version of 2.020. Refer to the SMAP Data Versions page for version information.
Note: The data product Science Version ID (example: Vv2020) consists of the first six characters of the data product Composite Release ID (CRID). The full CRID includes four additional digits that are to be found in individual granule metadata within the DataIdentifcation / DatasetIdentification / CompositeReleaseID field. These additional digits denote minor processing changes, such as runtime configuration and other minor changes that do not impact the science of the data product.</td></tr>
<tr><td>NNN</td><td>Number of times the file was generated under the same version for a particular date/time interval (002: 2nd time)</td></tr>
<tr><td>.[ext]</td><td>File extensions include:
<table>
<tr><td>.h5</td><td>HDF5 data file</td></tr>
<tr><td>.xml</td><td>XML Metadata file</td></tr>
</table></td></tr>
</table>

1.7 File Size

Each file is approximately 133 MB.

1.8 Volume

The daily data volume is approximately 133 MB.

1.9 Spatial Coverage

Coverage spans from 180°W to 180°E, and from approximately 85.044°N and 85.044°S.

1.9.1 Spatial Resolution

Level-4 carbon model inputs include the following spatial resolutions:

- 500 m resolution MODIS-based global PFT classification (from MCD12Q1 Type 5)
- 500 m Fraction of Photosynthetically Active Radiation (fPAR) data (from MOD15A2)
- 9 km resolution SMAP Level-4 soil moisture data (SPL4SMGP)
- ¼ degree pre-processed global, daily averaged meteorology data from the GEOS-5 Forward Processing (FP) system

Level-4 carbon model processing is conducted at 1 km EASE-Grid 2.0 resolution using spatially aggregated MODIS PFT and fPAR inputs. Level-4 carbon model daily global outputs are gridded using a 9 km EASE-Grid 2.0 projection consistent with the SMAP L4 soil moisture data used as input.

Note that while this product has a 9 km spatial resolution, it also retains sub-grid scale heterogeneity information as determined from the 1 km resolution processing using MODIS PFT and FPAR inputs.

For more details regarding inputs used in the carbon model, refer to the Data Sources section.

1.10 EASE-Grid 2.0

These data are provided on the global cylindrical EASE-Grid 2.0 (Brodzik et al. 2012). Each grid cell has a nominal area of approximately 9 x 9 km2 regardless of longitude and latitude.

EASE-Grid 2.0 has a flexible formulation. By adjusting a single scaling parameter, a family of multi-resolution grids that nest within one another can be generated. The nesting can be adjusted so that smaller grid cells can be tessellated to form larger grid cells. Figure 2 shows a schematic of the nesting to a resolution of 3 km (4872 rows x 11568 columns on global coverage), 9 km (1624 rows x 3856 columns on global coverage) and 36 km (406 rows x 964 columns on global coverage).

This feature of perfect nesting provides SMAP data products with a convenient common projection for both high-resolution radar observations and low-resolution radiometer observations, as well as for their derived geophysical products.

For more on EASE-Grid 2.0, refer to the EASE-Grid 2.0 Format Description.

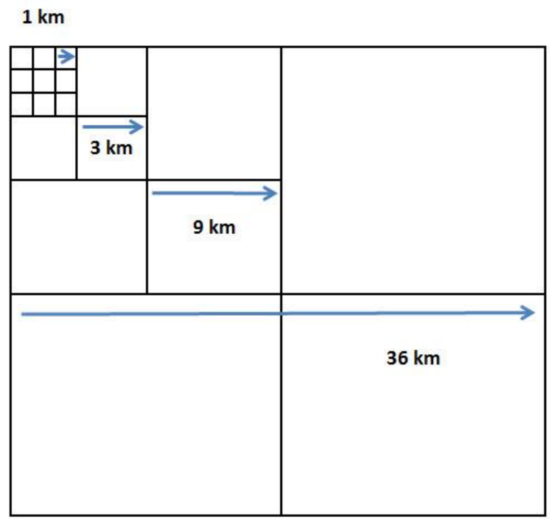

Figure 2. Perfect Nesting in EASE-Grid 2.0

1.11 Temporal Coverage

Coverage is continuous and spans from 31 March 2015 to 04 June 2018

1.11.1 Satellite and Processing Events

Due to instrument maneuvers, data downlink anomalies, data quality screening, and other factors, small gaps in the time series will occur. Refer to the SMAP On-Orbit Events List for Instrument Data Users page for details regarding these gaps.

SMAP On-Orbit Events List for Instrument Data Users
Master List of Bad and Missing Data

However, gaps in the SMAP time series do not affect this product. For the analytical variables, the ancillary MODIS fPAR 8-day climatology provides a fallback input source to help ensure there are no spatio-temporal gaps in the modeled data record.

For the period between 19 June and 23 July 2019, an extended gap occurred in the L1 - L3 SMAP products. During this period, the L4 Carbon data set was not informed by SMAP data. For more information on this SMAP outage, users should refer to the SMAP Post-Recovery Notice.

1.11.2 Latencies

FAQ: What are the latencies for SMAP radiometer data sets?

1.11.3 Temporal Resolution

Each Level-4 file is a daily composite. Calculations for this product are conducted at a daily time step in order to provide the necessary precision for resolving dynamic boreal vegetation phenology and carbon cycles (Kimball et al. 2009, Kim et al. 2012).

2 DATA ACQUISITION AND PROCESSING

This section has been adapted from Kimball et al. (2014), the ATBD for this product.

2.1 Background

Current capabilities for regional assessment and monitoring of NEE are limited by mismatches between bottom-up and top-down information sources. Atmospheric transport model inversions of CO_2 concentrations from sparsely distributed measurement stations provide information on seasonal patterns and trends in atmospheric CO_2 but little information on underlying processes; these methods are also too coarse to resolve carbon source-sink activity at scales finer than broad latitudinal and continental domains (Piao et al. 2007, Dargaville et al. 2002). Tower CO_2 flux measurement networks provide detailed information on stand-level NEE and associated biophysical processes, but little information regarding spatial variability in these processes over heterogeneous landscapes (Running et al. 1999). Estimates of NEE and component carbon fluxes from satellite remote sensing provide a means for scaling between relatively intensive stand-level measurement and modeling approaches, and top-down assessments from atmospheric model inversions.

To address these limitations, the primary objectives of the SPL4CMDL product are to:

- Determine NEE regional patterns and temporal behavior (daily, seasonal, and annual) to within the accuracy range of in situ tower measurement estimates of these processes;
- Link NEE estimates with component carbon fluxes (GPP and Rtot) and the primary environmental constraints to ecosystem productivity and respiration.

The SPL4CMDL algorithm supports carbon cycle science objectives by enabling detailed mapping and monitoring of spatial patterns and temporal dynamics of land-atmosphere CO_2 exchange, and the underlying carbon fluxes and environmental drivers of these processes. The SPL4CMDL product also links SMAP land parameter measurements to global terrestrial CO_2 exchange, including boreal ecosystems, reducing uncertainties about the 'missing sink' on land for atmospheric CO_2 .

Atmospheric transport model inversions of CO_2 concentrations indicate that the Northern Hemisphere terrestrial biosphere is responsible for much of the recent terrestrial sink strength for atmospheric carbon (Dargaville et al. 2002). Variability in land-atmosphere CO_2 exchange is strongly controlled by climatic fluctuations and disturbance, while uncertainty regarding the magnitude and stability of the sink are constrained by a lack of detailed knowledge on the response of underlying processes at regional scales (Denman et al. 2007, Houghton 2003).

The SPL4CMDL product enables quantification and mechanistic understanding of spatial and temporal variations in NEE over a global domain. NEE represents the primary measure of carbon (CO_2) exchange between the land and atmosphere, and the SPL4CMDL product is directly relevant to a range of applications including regional mapping and monitoring of terrestrial carbon stocks and fluxes, climate and drought related impacts on vegetation productivity, and atmospheric transport model inversions of terrestrial source-sink activity for atmospheric CO_2.

For more background information, refer to Section 2.3: Historical Perspective in the ATBD for this product (Kimball et al. 2014).

2.2 Acquisition

The following data sources are used as input to calculating this Level-4 carbon product:

- SMAP L4 9 km EASE-Grid Surface and Root Zone Soil Moisture Geophysical Data, Version 4 (SPL4SMGP)
- GMAO GEOS-5 Forward Processing (FP) Model Data: Daily surface meteorology from observation-corrected global atmospheric model analysis
- NASA EOS Terra MODIS fPAR 8-day Data, Version 6 (MOD15A2): Canopy fPAR and land cover classification; if MOD15A2 data are unavailable, the following back-up sources are used to calculate fPAR:
    - SMAP L4 Carbon Model ancillary MODIS fPAR 8-day Climatology: Primary back-up source
    - NASA EOS Aqua MODIS fPAR 8-day Data, Version 6 (MYD15A2): Canopy fPAR and land cover classification: Secondary back-up source
    - VIIRs NDVI Data (VVI3P): Secondary back-up source

In addition, ancillary data sources used as input to calculating this Level-4 carbon product are listed in Table 3.

Table 3. Primary Ancillary Inputs to the SPL4CMDL Algorithm

Parameter	Units	Type	Spatial Resolution	Source
fPAR	%	Dynamic (8-day)	1 km	MODIS (MOD15A2*)
R_{sw}	MJ m - 2 d - 1	Dynamic (daily)	9 km**	GEOS-5
T_{mn}	°C	Dynamic (daily)	9 km**	GEOS-5
VPD	Pa	Dynamic (daily)	9 km**	GEOS-5
SM	% Sat.	Dynamic (daily)	9 km	SPL4SMGP
SM_{rz}	% Sat.	Dynamic (daily)	9 km	SPL4SMGP
T_s	°C	Dynamic (daily)	9 km	SPL4SMGP
F/T	Discrete class	Dynamic (daily)	9 km**	GEOS-5***
Land Cover Class	Discrete class	Static	1 km	MODIS (MOD12Q1)
fPAR Climatology	%	Static (8-day)	1 km	MODIS (MOD15A2)
Additional Inputs for Algorithm Options				
VI (NDVI)	Dimensionless	Dynamic (8-day)	1 km	MODIS (MOD13A2, MYD13A2), VIIRS (VVI3P)
Recovery Status	Years	Static	1 km	MODIS (MOD13A2, MYD13A2)

* MOD indicates data acquired by the MODIS instrument on the Terra satellite; MYD indicates data acquired by the MODIS instrument on the Aqua satellite.

** The native resolution of GEOS-5 FP fields is ¼ degree (latitude) by 3/8 degree (longitude); SPL4CMDL processing internally resamples these fields to 9 km.

*** Due to the loss of the SMAP radar instrument and operational freeze/thaw (F/T) classification product, SPL4CMDL uses the GMAO GEOS-5-modeled TSURF parameter to define F/T conditions in the carbon model.

**** Derived from finer scale (500 m resolution) MODIS data records and spatially aggregated to 1 km resolution for carbon model processing.

2.3 Baseline Alogrithm

The baseline SPL4CMDL algorithm uses daily inputs from the SMAP Level-4 soil moisture stream to define soil moisture and frozen temperature constraints to vegetation productivity, ecosystem respiration, and NEE. The algorithm provides estimates of NEE (g C m-2 day-1) and component carbon fluxes for global vegetated land areas at mean daily intervals; the product defines sub-grid scale mean and variability in carbon fluxes for dominant and sub-dominant vegetation classes within each grid cell as determined from finer scale ancillary land cover classification and fPAR inputs. The target accuracy for the SPL4CMDL product is to attain a mean annual unbiased RMSE (ubRMSE) accuracy for NEE within 30 g C m-2 yr-1 or 1.6 g C m-2 day-1, commensurate with the estimated accuracy of in situ tower measurements (Baldocchi et al. 2008, Richardson 2005, Richardson 2008). The baseline 1 km SPL4CMDL spatial resolution is similar to the sampling footprint of CO2 flux measurements from the global tower network (Running et al. 1999, Baldocchi et al. 2008). Secondary products of scientific value produced during SPL4CMDL processing include surface (<10 cm depth) Soil Organic Carbon (SOC) stocks (g C m-2), vegetation Gross Primary Production (GPP), heterotrophic soil and litter respiration (Rh), dimensionless (0-100 percent) environmental constraint indices for GPP and Rh, and detailed data Quality Assessment (QA) metrics for NEE.

The SPL4CMDL algorithm consists of Light Use Efficiency (LUE) and terrestrial carbon flux model components used to estimate GPP, respiration, residual NEE carbon fluxes, and underlying SOC pools on a daily basis. The baseline SPL4CMDL algorithm is summarized in Figures 4a and 4b for respective LUE and carbon flux model components. The approach has structural elements similar to the Century (Parton et al. 1987, Ise and Moorcroft 2006) and CASA (Potter et al. 1993) soil decomposition models and the operational MOD17 GPP algorithm (Zhao et al. 2005, Zhao and Running 2010), but is adapted for use with daily biophysical inputs derived from both global satellite and model analysis data (Kimball et al. 2009, Yi et al. 2013). The current SPL4CMDL algorithm baseline was developed from earlier versions and pre-launch development and testing, and incorporates recommendations from external SPL4CMDL algorithm reviews (for example, Kimball et al. 2009).

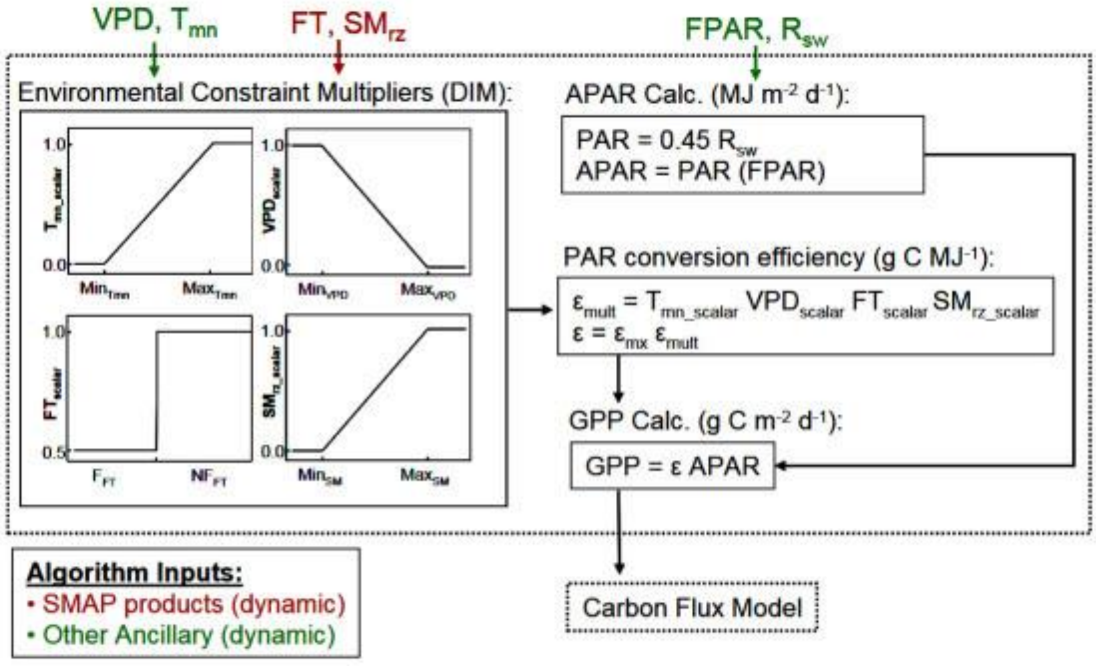

Figure 3a. Baseline Light Use Efficiency (LUE) Carbon Model Structure for Estimating GPP. Arrows denote the primary pathways of data flow, while boxes denote the major process calculations. Primary inputs include daily root zone soil moisture (SMrz) and landscape freeze/thaw (FT) status from SMAP Level-4 soil moisture products (in red), and other dynamic ancillary inputs (in green) including MODIS (MOD15) fPAR and GMAO GEOS-5 daily surface meteorology, including vapor pressure deficit (VPD), minimum air temperature (Tmn) and incident solar shortwave radiation (Rsw). Model calculations are performed at 1 km spatial resolution using dominant vegetation class and BPLUT response characteristics for each grid cell defined from a global land cover classification. The resulting GPP calculation is a primary input to the Level-4 carbon terrestrial carbon flux model below (Figure 4b).

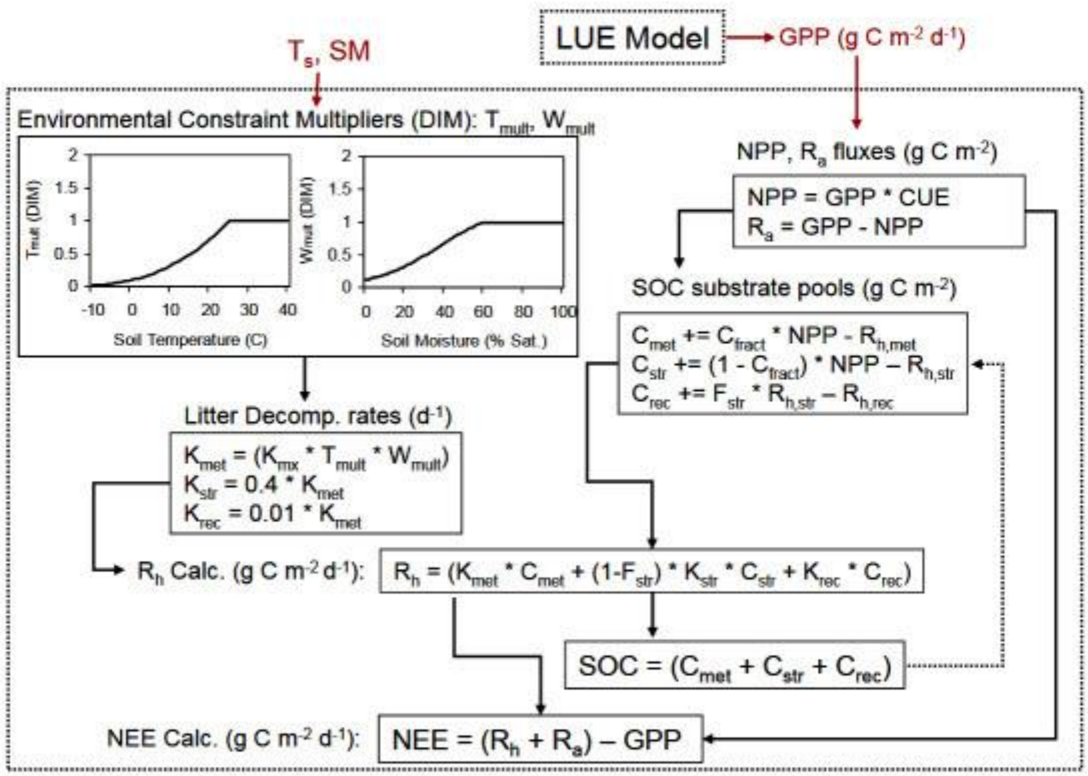

Figure 3b. Terrestrial Carbon Flux Model for Estimating NEE. Primary algorithm inputs (in red) include daily GPP from the LUE model (Figure 4a), and surface soil moisture (SM) and surface temperature (Ts) from the SMAP Level-4 product. NEE is the primary (validated) output, while GPP, respiration (Rh + Ra), and SOC are secondary (research) outputs.

Dynamic daily inputs to the SPL4CMDL algorithms include satellite optical-IR remote sensing MODIS-based fPAR, GEOS-5 surface meteorology (Rsw, Tmn, VPD) and associated SPL4SMGP soil moisture (SMrz) which provide primary inputs to a LUE algorithm to determine GPP, where Rsw is incoming shortwave solar radiation (MJ m^{-2} d^{-1}); Tmn is minimum daily 2-m air temperature (°C), VPD is atmosphere vapor pressure deficit (Pa), and SMrz is the integrated surface to root zone (0-1m depth) soil moisture (% Sat.). The SPL4CMDL dynamic inputs also include GEOS-5 freeze/thaw classification defined byÂ surface temperature (Ts, °C) or the SMAP sensor constraints to productivity and autotrophic respiration calculations. SMAP Level-4 surface soil moisture (≤ 5 cm depth) and soil temperature are used as primary drivers of the soil decomposition and Rh calculations. Static inputs to the SPL4CMDL algorithms include a global land cover classification, which is used to define the major plant functional types and associated biome-specific Biome Properties Look-Up Table (BPLUT) response characteristics for each vegetated grid cell within the product domain. The BPLUT parameters are defined for up to 8 global vegetation (PFT) classes; the model parameters for each global PFT class were calibrated prior to SPL4CMDL operational production by optimizing carbon model NEE calculations against tower eddy covariance measurement based daily NEE observations from global FLUXNET sites representing the major PFT classes (Baldocchi 2008). The land cover classification used for SPL4CMDL processing is consistent with the one used in the production of the fPAR inputs. All model inputs are available as satellite remote sensing derived products or from model (GEOS-5) analysis.

The resulting SPL4CMDL parameters enable characterization of spatial patterns and daily temporal fidelity in NEE, underlying carbon fluxes and SOC pools, and their primary environmental drivers. The resulting fine scale (1 km resolution) SPL4CMDL outputs are spatially aggregated to the coarser 9 km resolution final product grid by weighted linear averaging of outputs according to the fractional cover of individual PFT classes represented within each 9 km grid cell and defined by the underlying 1 km resolution MODIS PFT map; the sub-grid scale means from individual PFT classes are preserved for each 9 km grid cell, while proportional vegetation cover information is included in the product metadata, allowing the coarse resolution data to be decomposed into the relative contributions from individual PFT classes within each cell. These outputs are designed to facilitate improved algorithm and product accuracy over heterogeneous land cover areas, and product outputs that are more consistent with the mean sampling footprint of most tower CO2 flux measurement sites (Baldocchi 2008, Chen et al. 2012).

2.3.1 Algorithm Options

For more information regarding algorithm options, refer to the ATBD, for this product.

2.3.2 Ancillary Data

Ancillary data required as input for the algorithms are summarized in Table 4. For in-depth information on ancillary data, refer to the ATBD, Section 3.2: Ancillary Data Requirements.

For more information regarding the algorithm, refer to the ATBD for this product.

2.4 Processing Steps

Written by the University of Montana's Numerical Terradynamic Simulation Group (NTSG), the SPL4CMDL science code was transferred from NTSG to the NASA Global Modelling and Assimilation Office (GMAO) for translation and implementation as operational code in conjunction with SMAP Level-4 soil moisture production within the GMAO Level-4 SMAP Science Data Processing System (SDS).

To generate the SPL4CMDL product, the processing software:

1. Ingests SPL4SMGP daily files, MODIS-derived 8-day fPAR files, and GEOS-5 daily surface meteorology data.

2. The ingested data are then inspected for retrievability criteria according to input data quality, ancillary data availability, and land cover conditions.
3. Two pre-processor codes, one for fPAR data and one for global meteorology data, are then executed each day to temporally aggregate and resample these respective inputs for use by the baseline algorithm software. When retrievability criteria are met, the production software invokes the baseline retrieval algorithm to generate the daily carbon model outputs.

SPL4CMDL calculations are conducted at 1 km resolution, benefiting from finer scale (500 m) MODIS fPAR and land cover inputs. The simulations have also been conducted in a consistent global EASE-Grid 2.0 projection format. Model simulations for each 1 km grid cell are conducted using the corresponding (nearest-neighbor) 9 km resolution SMAP Level-4 Soil Moisture and GEOS-5 inputs. The MODIS (MOD/MYD15) fPAR product is produced at 500 m resolution and 8-day temporal fidelity from both NASA EOS Terra and Aqua sensor records.

MODIS fPAR operational products are obtained in a tile-based sinusoidal projection. Preprocessing of these data prior to the SPL4CMDL ingestion involves reprojecting from sinusoidal to 1 km resolution global cylindrical EASE-Grid projection formats, followed by trailing nearest-neighbor temporal interpolation of MOD15A2 good Quality Control (QC; relatively cloud-free with favorable surface conditions) 8-day fPAR series to each daily time step. Missing or low QC 8-day fPAR data are gap filled on a grid cell-wise basis using an ancillary fPAR mean 8-day climatology constructed from the long-term (10+ year) MODIS record. The resulting fPAR data are combined with daily biophysical inputs from GEOS-5 and SPL4SMGP data to estimate NEE, component carbon fluxes (GPP and Rh) and surface SOC pools. SPL4CMDL computes daily Environmental Constraint (EC) indices which influence the GPP and NEE flux calculations, including the estimated bulk environmental reduction to PAR conversion efficiency (εmult), low soil moisture and temperature constraints (Wmult, Tmult) to soil decomposition and Rh calculations, and freeze/thaw (F/T) status within each 9 km grid cell. These environmental constraint indices are provided in SPL4CMDL files as the EC/emult_mean, EC/wmult_mean, EC/tmult_mean and EC/frozen_area respective data fields.

2.5 Quality, Errors, and Limitations

2.5.1 Error Sources

Many sources of error contribute to the uncertainty in the SPL4CMDL product. The key sources of error or uncertainty to the SPL4CMDL algorithm are:

1. Errors in the ancillary 8 day fPAR inputs

2. Errors in the SPL4SM soil moisture and temperature inputs

3. Errors in the GEOS-5 daily surface meteorology inputs

4. Uncertainty in the internal model parameterization, initialization, and calibration parameters

For more information about error sources refer to the ATBD for this product.

2.5.2 Quality Assessment

For in-depth details regarding the quality of these Version 2 data, refer to the following reports:
Validated Assessment Report
Beta Assessment Report

2.5.2.1 Quality Overview

SMAP products provide multiple means to assess quality. Each product contains bit flags, uncertainty measures, and file-level metadata that provide quality information. For information regarding the specific bit flags, uncertainty measures, and file-level metadata contained in this product, refer to the Product Specification Document.

Each HDF5 file contains metadata with Quality Assessment (QA) metadata flags that are set by the GMAO prior to delivery to National Snow and Ice Data Center Distributed Active Archive Center (NSIDC DAAC). A separate metadata file with an .xml file extension is also delivered to NSIDC DAAC with the HDF5 file; it contains the same information as the HDF5 file-level metadata.

2.5.2.2 Quality Flags

Quality Assessment (QA) fields are also provided with metadata from MODIS fPAR and SPL4SM inputs to the SPL4CMDL algorithms. These QA fields incorporate expected model uncertainty propagating from input driver uncertainty including SPL4SMGP, GEOS-5 FP, and MODIS fPAR. This QA input error information was assigned by comparing unbiased Root Mean Square Errors (ubRMSE) relative to global historical flux tower benchmark data during SPL4CMDL pre-launch calibration. Input errors are propagated during SPL4CMDL 1 km model calculations using standard error propagation procedures employing the SPL4CMDL model Jacobian and simplifying independence assumptions. Resulting 1 km NEE ubRMSE fields are quadratically averaged to 9 km output fields for each PFT class as defined from 1 km MOD12Q1 land cover and then posted as the NEE QA ubRMSE geophysical variable (g C m^{-2} d^{-1}). The resulting QA information has been evaluated and refined through post-launch SPL4CMDL Cal/Val activities using concurrent eddy covariance CO2 flux measurements from global tower measurement networks (Baldocchi 2008), comparisons with other similar global carbon products, and algorithm sensitivity studies over the observed range of environmental variability. The above-described QA fields are provided in SPL4CMDL files as the QA/nee_rmse_mean and QA/nee_rmse_pft{1..8}_mean fields. Refer to the SPL4CMDL Product Specification Document (PSD) Version 2.0 for additional details.

Quality control bit flags are provided in SPL4CMDL files to identify retrieval conditions including use of alternative ancillary data sets and exceedance of expected output field value ranges. Alternative ancillary conditions indicated in the QC bit flags include the use of alternative fPAR sources in place of baseline MODIS (MOD15) fPAR inputs, potential gaps in the SPL3SMA input stream, and instances where the ancillary fPAR 8-day climatology is used in place of the dynamic best QC MODIS fPAR input stream to estimate GPP. Expected PFT class specific range thresholds for each state variable (NEE, GPP, Rh, and SOC) have been established from dynamic algorithm simulations using long-term (10+ year) daily data input records from pre-launch data sources similar to those used for post-launch SPL4CMDL production, including MODIS (MOD15) fPAR, freeze-thaw status (Kim et al. 2012), and MERRA surface meteorology (Yi et al. 2011). These post-launch diagnostics are provided in SPL4CMDL files in the QA/carbon_model_bitflag data field for additional user evaluation. Table 7 indicates the bit-field positions for the above-described flags. A copy of Table 7 is also provided within each SDS file as metadata for quick reference; refer to the QA/carbon_model_bitflag data field.

Table 4. QC Bit Flag Fields, Names, Positions, and Description Metadata

Bit Flag Name	**Bit Positions {Start, End}**	**Number of Bits**	**Value Range**	**Description**
NEE bit	00 – 00	1	{0|1}	0 = NEE within valid range; 1 = out of valid range
GPP bit	01 – 01	1	{0|1}	0 = GPP within valid range; 1 = out of valid range
Rh bit	02 – 02	1	{0|1}	0 = Rh within valid range; 1 = out of valid range
SOC bit	03 – 03	1	{0|1}	0 = SOC within valid range; 1 = out of valid range
PFT dominant	04 – 07	4	{1..8}	Most frequently occurring (dominant) vegetated PFT class as defined from qa_count
QA score	08 – 11	4	{0,1,2,3}	Relative nee_mean error as ranked by nee_rmse_mean: 0 = (RMSE<1 g C m^{-2} d^{-1}); 1 = (1<=RMSE<2 g C m^{-2} d^{-1}); 2= (2=<RMSE<3 g C m^{-2} d^{-1}); 3 = (RMSE> = 3 g C m^{-2} d^{-1})
GPP method	12 – 12	1	{0|1}	0 = derived GPP using 8-day FPAR or NDVI input, 1 = derived GPP via FPAR or NDVI climatology
NDVI method	13 – 13	1	{0|1}	0 = derived GPP using FPAR; 1 = derived GPP using NDVI
F/T method	14 – 14	1	{0|1}	0 = used SPL3SMA F/T; 1 = used GEOS-5 surface temperature

Bit Flag Name	Bit Positions {Start, End}	Number of Bits	Value Range	Description
IsFill*	15 – 15	1	{0|1}	0 = is NOT fill value (simulation performed for one or more 1 km grid cells within 9 km grid cell), 1 = is fill value (no 1 km simulation performed within 9 km grid cell). Fill values occur for non-land, non-vegetated, and/or grid cells otherwise lacking valid FPAR data record.
* When IsFill = 1, then all other bit fields will have value 1 and the entire uint16 integer will evaluate to 65534. Users should therefore check the value of IsFill prior to referencing other bit fields.				

For more information, such as algorithm testing procedures and Calibration and Validation (Cal/Val) activities, refer to the ATBD. For more information regarding data flags, refer to the Product Specification Document.

3 INSTRUMENTATION

3.1 Description

For a detailed description of the SMAP instrument, visit the SMAP Instrument page at the JPL SMAP website.

4 SOFTWARE AND TOOLS

For tools that work with SMAP data, refer to the Tools web page.

5 VERSION HISTORY

Table 5. Summary of Version Changes

Version	Date	Version Changes
Version 1	October 2015	First public data release
Version 2	April 2016	Changes to this version include: Transitioned to Validated-Stage 2 Using SPL4SMAU V2 Validated and SPL4SMGP V2 Validated data as input Update to process radiometer data from 2015-03-31 to 2015-04-12 Some data fields renamed from *_av to *_mean Updated to have have continuous RMSE-based "quality" fields instead of the categorical quality flag in V1
Version 3	July 2017	Changes to this version include: Uses dynamic 8-day fPAR inputs obtained from the latest (Collection 6) MODIS fPAR record at 500 m resolution. The preprocessor was updated to handle the finer resolution MODIS Collection 6 inputs, which are interpolated to 1 km resolution prior to model processing. The prior (Version 2) processor used MODIS Collection 5 fPAR inputs, which were derived at 1 km resolution. Updated the ancillary MODIS fPAR 8-day climatology used for fPAR gap-filling as an L4C model preprocessing step to reflect new MODIS Collection 6 fPAR inputs. The fPAR climatology is derived from a longer 14-year (2000-2014) MODIS record relative to the original 12-year (2000-2012) Collection 5 fPAR record used in Version 2 processing. For each grid cell, a sine-curve-based seasonal fPAR climatology curve is now used to identify and screen anomalous 8-day fPAR variations in the preprocessor. This change reduces impacts of anomalous fPAR temporal variations that may not be captured by the MODIS fPAR product quality control (QC) flags, particularly during seasonal transitions at northern latitudes. Updated and recalibrated the ancillary Biome Properties Look-Up Table (BPLUT) and re-initialized the model initial global soil organic carbon (SOC) pools to reflect new MODIS Collection 6 fPAR inputs. The BPLUT calibration was conducted using global historical FLUXNET in situ tower eddy covariance CO2 flux measurement records for representative global land cover types using a similar step-wise calibration procedure employed for the Version 2 product. A minor bug fix to the post-processor was made to ensure that all grid cell no-data fill values are identified with a consistent -9999 notation; the prior Version 2 product erroneously assigned some no-data values as -999900.

Version	Date	Version Changes
V3	January 2021	Changes to this version include: Extended temporal coverage to 04 June 2018
V4	June 2018	Changes to this version largely affect model inputs and ancillary files rather than changes to the internal model structure or code. Note that Version 4 is slightly better than Version 3 in RMSE terms, with improvement generally larger for drier sites. Specific changes include: The carbon model biome properties lookup table (BPLUT) has been calibrated using an augmented FLUXNET global tower site record which includes more calibration sites (335 sites compared to 228 sites for V3), expanded tower data records extending to at least 2015, and the addition of new tower sites representing more land cover types. Revised Level-4 carbon global model calibration and SOC initialization using an extended (2000-2017) MODIS fPAR (V006) record and the latest SMAP Nature Run (NRv7.2) climate data records. Implemented minor changes to spatial weighting of calibration tower sites within a model grid cell and reduced the outlier influence on model response curve fitting.

6 RELATED DATA SETS

SMAP Data at NSIDC | Overview

SMAP Radar Data at the ASF DAAC

7 RELATED WEBSITES

SMAP at NASA JPL

8 CONTACTS AND ACKNOWLEDGMENTS

Investigators

John Kimball, Lucas Jones, Joe Glassy

Numerical Terradynamic Simulation Group (NTSG)

College of Forestry & Conservation

The University of Montana

Missoula, MT 59812-1049 USA

Rolf Reichle

NASA Goddard Space Flight Center

Global Modeling and Assimilation Office

Mail Code 610.1

8800 Greenbelt Rd
Greenbelt, MD 20771 USA

9 REFERENCES

Brodzik, M. J., B. Billingsley, T. Haran, B. Raup, and M. H. Savoie. 2012. EASE-Grid 2.0: Incremental but Significant Improvements for Earth-Gridded Data Sets. ISPRS Int. J. Geo-Inf. 1(1):32-45. http://dx.doi.org/10.3390/ijgi1010032.

Brodzik, M. J., B. Billingsley, T. Haran, B. Raup, and M. H. Savoie. 2014. Correction: Brodzik, M. J. et al. EASE-Grid 2.0: Incremental but Significant Improvements for Earth-Gridded Data Sets. ISPRS Int. J. Geo-Inf 2012. 1(1):32-45 ISPRS Int. J. Geo-Inf. 3(3):1154-1156. http://dx.doi.org/10.3390/ijgi3031154.

Dargaville, R. A.D. McGuire, and P. Rayner. 2002. Estimates of large-scale fluxes in high latitudes from terrestrial biosphere models and an inversion of atmospheric CO2 measurements. Climatic Change 55:273–285.

Entekhabi, D. et al. 2014. SMAP Handbook–Soil Moisture Active Passive: Mapping Soil Moisture and Freeze/Thaw from Space. Pasadena, CA USA: SMAP Project, JPL CL#14-2285, Jet Propulsion Laboratory.

Glassy, J., J. S. Kimball, L. Jones, R. H., Reichle, R. A. Lucchesi, J. V. Ardizzone, G.-K. Kim, and B. H. Weiss. 2016. Soil Moisture Active Passive (SMAP) Mission Level 4 Carbon (L4_C) Product Specification Document. GMAO Office Note No. 11, Version 2.0. NASA Goddard Space Flight Center, Greenbelt, MD, USA. (SMAP_L4C_PSD_v2_0_20160421.pdf, 1.3 MB)

Kimball, J. S., L. A. Jones, J. Glassy, E. N. Stavros, N. Madani, R. H. Reichle, T. Jackson, and A. Colliander. 2016. Soil Moisture Active Passive (SMAP) Mission Assessment Report for the Version 2 Validated Release L4_C Data Product. GMAO Office Note No. 13 (Version 1.0), 37 pp, NASA Goddard Space Flight Center, Greenbelt, MD, USA. (SMAP-Validation_Report_L4_C_20160421.pdf, 2.25 MB)

Kimball, J. S., L. A. Jones, J. Glassy, E. N. Stavros, N. Madani, R. H. Reichle, T. Jackson, and A. Colliander. 2015. Soil Moisture Active Passive (SMAP) Project Calibration and Validation for the L4_C Beta-Release Data Product. NASA/TM–2015-104606, Vol. 42, 37 pp. (Kimball806.pdf, 13.7 MB)

Kimball, J. S., L. A. Jones, J. P. Glassy, and R. Reichle. 2014. SMAP Algorithm Theoretical Basis Document, Release A: L4 Carbon Product. SMAP Project, JPL D-66484, Jet Propulsion Laboratory, Pasadena CA. 76 pp. (271_L4_C_RevA_web.pdf, 2.7 MB)

National Research Council (NRC). 2007. Earth Science and Applications from Space: National Imperatives for the Next Decade and Beyond (Executive Summary). National Academy of

Sciences, National Academies Press, Washington DC. 35pp.
http://www.nap.edu/catalog/11820.html.

Yi, Y., J. S. Kimball, L. A. Jones, R. H. Reichle, R. Nemani, and H. A. Margolis. 2013. Recent climate and fire disturbance impacts on boreal and arctic ecosystem productivity estimated using a satellite-based terrestrial carbon flux model. J. Geophys. Res. Biogeosci. 118, 1-17.

10 DOCUMENT INFORMATION

10.1 Publication Date

7 January 2019

10.2 Date Last Updated

28 January 2021