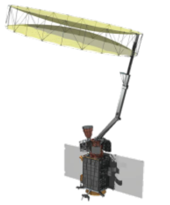

Soil Moisture Active Passive (SMAP)

Ancillary Data Report

Static Water Fraction

Preliminary, v.2
SMAP Science Document no. 045

Steven Chan
Jet Propulsion Laboratory
California Institute of Technology
Pasadena, CA

January 21, 2013
JPL D-53059

Jet Propulsion Laboratory
California Institute of Technology

Preface

The SMAP Ancillary Data Reports provide descriptions of ancillary data sets used with the science algorithm software in generation of the SMAP science data products. The Ancillary Data Reports may undergo additional updates as new ancillary data sets or processing methods become available. The most recent versions of the ancillary data reports will be made available, along with the Algorithm Theoretical Basis Documents (ATBDs), at the SMAP web site http://smap.jpl.nasa.gov/science/dataproducts/ATBD/.

Table of Contents

1 Overview

1.1 Purpose

The purpose of this report is to develop a static water fraction dataset for use in generating SMAP science data products. The static water fraction dataset is one of a suite of ancillary datasets used by the SMAP science processing algorithms. The algorithms and ancillary data are described in SMAP algorithm theoretical basis documents (ATBDs) and ancillary data reports. The ATBDs and ancillary data reports are listed in Appendices A and B and are available at the SMAP web site: http://smap.jpl.nasa.gov/science/dataproducts/ATBD/.

1.2 Requirement

Water fraction refers to the fraction (by area) of water relative to land in a given spatial extent. Depending on the context of processing, this spatial extent may refer to the field-of-view (FOV) of the antenna, or simply a grid cell of a given grid resolution. In all cases, it varies between 0 (no water) and 1 (all water).

The static water fraction dataset will be used to discriminate between land and water pixels, and to provide information for brightness temperature correction for pixels containing partial land and water. As the SMAP radar-based water detection algorithm matures, information regarding the presence of transient water can be used to augment the static water information and serve the two purposes above more accurately. The focus of this report is on the development of a global water fraction dataset based on static water information only.

2 Dataset Description and Selection

2.1 Source Datasets

There are a number of candidate data sources available for deriving global static water fraction. Below is a list of primary sources considered here. Given the large number of data sources, it is not practical to fully evaluate all of them. The investigation is therefore limited to the more recent data sources and those with a broader known user base.

2.2 Dataset Selection

Among the data sources listed in Table 1, the MOD44W dataset is the one considered most appropriate for deriving a global static water fraction dataset for SMAP. Formally, it is an official product developed and released by the MODIS Land Science Team in 2009 to fulfill "*the current needs of the terrestrial remote sensing community working at 250m spatial resolution*" (Carroll et al., 2009). The product uses a radar-based Shuttle Radar Topography Mission (SRTM) Water Body Dataset (SWBD, 2012) as the baseline data source between 56S and 60N, and supplements it with MODIS 250-m MOD44C product wherever SWBD has quality or availability issues. For regions beyond 60N and 56S, the MODIS 250-m products are used (NASA, 2007; Haran, 2006).

Table 1: Candidate data sources available for deriving a global static water fraction database

Dataset	Description	Scale	Notes
MODIS-IGBP	Terra/Aqua MODIS land cover (V005) in IGBP classes	500m	Annual, 2001-2007
MODIS-UMD	Terra/Aqua MODIS land cover (V005) in UM classes	500m	Annual, 2001-2007
MODIS-NPP	Terra/Aqua MODIS land cover (V005) in NPP classes	500m	Annual, 2001-2007
MODIS-LAI/fPAR	Terra/Aqua MODIS land cover (V005) in LAI/fPAR classes	500m	Annual, 2001-2007
MODIS PFT	Terra/Aqua MODIS land cover (V005) in Plant Functional Type classes	500m	Annual, 2001-2007
AVHRR-IGBP	AVHRR land cover in IGBP classes	1000m	One time, 1992-1993
AVHRR-UM	AVHRR land cover in UM classes	1000m	One time, 1981-1994
ECOCLIMAP	MODIS, CORINE, PELCOM land cover in IGBP and UM classes	1000m	One time, 1992-1993
SRTM	Shuttle Radar Topography Mission	30m	One time, 2000. 56S-60N
GLDW	Level 3 Global Lakes and Wetland Database	1000m	One time
MOD44W	Terra/Aqua MODIS land cover (V005) merged with SRTM	250m	One time

MOD44W represents the MODIS Land Science Team's most recent effort in water/land mask development at the highest spatial resolution (250 m). In the process, the team surveyed a number of existing data sources and addressed their shortcomings according to the combined use of radar and optical data. The end product is a high-quality (in terms of grid resolution and spatial continuity) global land/water mask utilizing the best data sources available. Given the many improvements of MOD44W over existing data sources, the MODIS Land Science Team endorsed the use of MOD44W as the standard water mask in upcoming MODIS Collection 6 (currently 5) reprocessing. Many of the data sources identified by SMAP in Table 1 as potential data sources had earlier been evaluated by the MODIS Land Science Team. Table 2 provides further details of MOD44W.

3 Processing

3.1 Data Acquisition and Processing

Data were acquired using the NASA WIST system (now upgraded and renamed as REVERB) (REVERB, 2012). The global coverage of MOD44W is represented by 318 10° × 10° "tiles" for a total data volume of about ~76 MB. Each tile contains two 4800-by-4800 data arrays (one for data and the other for QA) stored as HDF-EOS Grid objects. As Grid objects, these tiles do not contain explicit geolocation data. To populate the actual latitude/longitude coordinates two additional tools (HDF-EOS2 Dumper and GDconvert_ij2ll.c) are required.

Table 2: Specifications of the MOD44 dataset.

MOD44W Specifications	Description
Developers	Mark Carroll, Charlene DiMiceli, John Townshend, Praveen Noojipady, Robert Sohlberg
Spatial resolution	Global coverage at 250 meters
Temporal resolution	Static
Data sources	1. 60S-60N: SRTM Water Body Dataset, supplemented with supplemented with MODIS 250-m MOD44C wherever needed 2. 60N-90N: MODIS 250-m MOD44C (16-day composites from 8 yrs/Terra and 6 yrs/Aqua) 3. 60S-90S: MODIS 250-m Mosaic of Antarctica
Improvements over existing data sources	1. Relative to SRTM: Expanded geographical coverage and reduced discontinuities in major river networks 2. Relative to MODIS: Improved spatial resolution and cloud-covered areas resolved by SRTM 3. Relative to GLWD: Improved spatial resolution, reduced geolocation errors, and updated coastline database
Raw data format	10° × 10° HDF-EOS tiles on sinusoidal projection
Availability	USGS DAAC
Documentation	UMD Global 250 meter Land Water Mask User Guide
Contact	Dr. Mark Carroll. Email: markc@umd.edu
URL	http://www.landcover.org/data/watermask/

Once the latitude/longitude coordinates are available the data are processed such that the numbers of 1's (indicating water pixels) and 0's (indicating land pixels) at 250-m resolution are tracked and used to derive static water fraction maps at coarser grid (EASE grid) resolutions. The exact steps are as follows:

1. For each 250-m grid cell containing either water (1) or land (0), convert its latitude/longitude coordinates into the EASE-grid row and column indices.
2. Repeat the above procedure for the entire global array, while doing so:
   a. Keep count ***N*** of water pixels falling into each EASE-grid cell
   b. Keep count ***M*** of land pixels falling into each EASE-grid cell
3. Compute the fraction $N / (N + M)$ for each EASE-grid cell
4. Repeat Steps 1 through 3 for the 1, 3, 9 and 36 km EASE-grid projections.

The result of the above processing steps is a set of static water fraction maps on 1-, 3-, 9-, and 36-km EASE-Grid projections.

3.2 Data Accuracy

Much of the information was extracted from the Validation section in Carroll et al (2009) where it was stated that "the new 250-m water mask (MOD44W) is a dramatic improvement over the current 1km raster mask that is used in MODIS data processing and many other purposes. The product will be included in the MODIS Collection 6 reprocessing as the standard water mask used in the creation of many of the MODIS standard products. It will also be incorporated into the MODIS Vegetation Continuous Fields product as well as the MODIS Land Cover product."

- On the use of the SWBD product between 56S and 60N: "… The absolute vertical accuracy was determined to be ~9m and the absolute geolocation accuracy was determined to be ~8m." (Rodriguez et al., 2006).
- On the use of the MOD44C product between 60N and 90N: The resulting MOD44W product shows a 2% commission error relative to the National Land Cover Dataset (NLCD) product, whose accuracy was stated to be 90% (Homer et al., 2004).
- On the use of the Mosaic of Antarctica product between 56S and 90S: The developers "found no discrepancies greater than 125 m for fixed objects in well mapped areas in more than 260 scenes." (Haran, 2006).

3.3 Results

Figures 1-3 show global static water fraction maps at different grid resolutions on the EASE-Grid 2.0 projection.

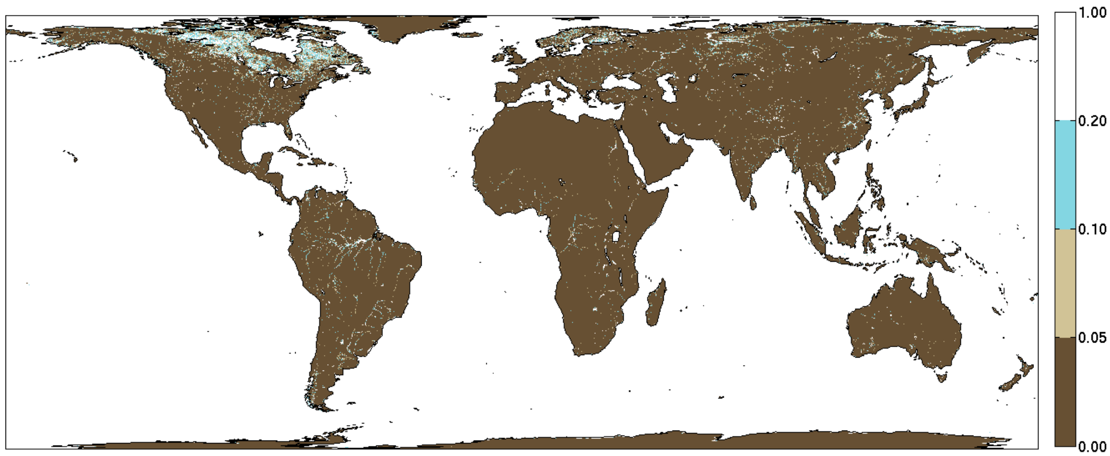

Figure 1. Global static water fraction at 3 km

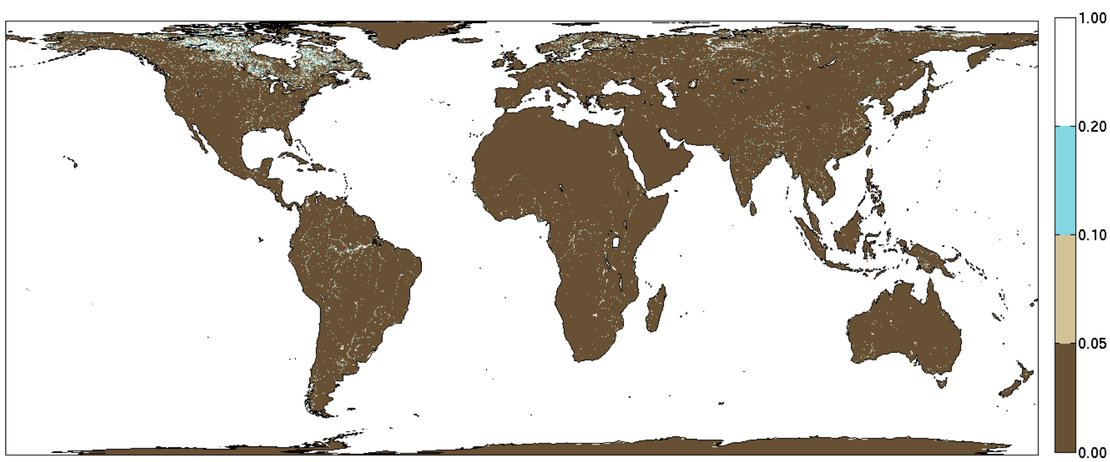

Figure 2. Global static water fraction at 9 km

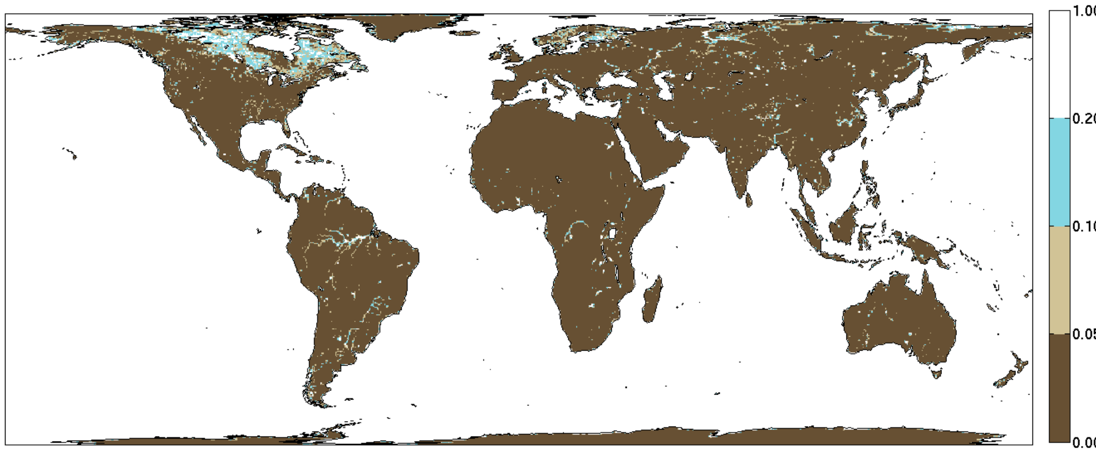

Figure 3. Global static water fraction at 36 km

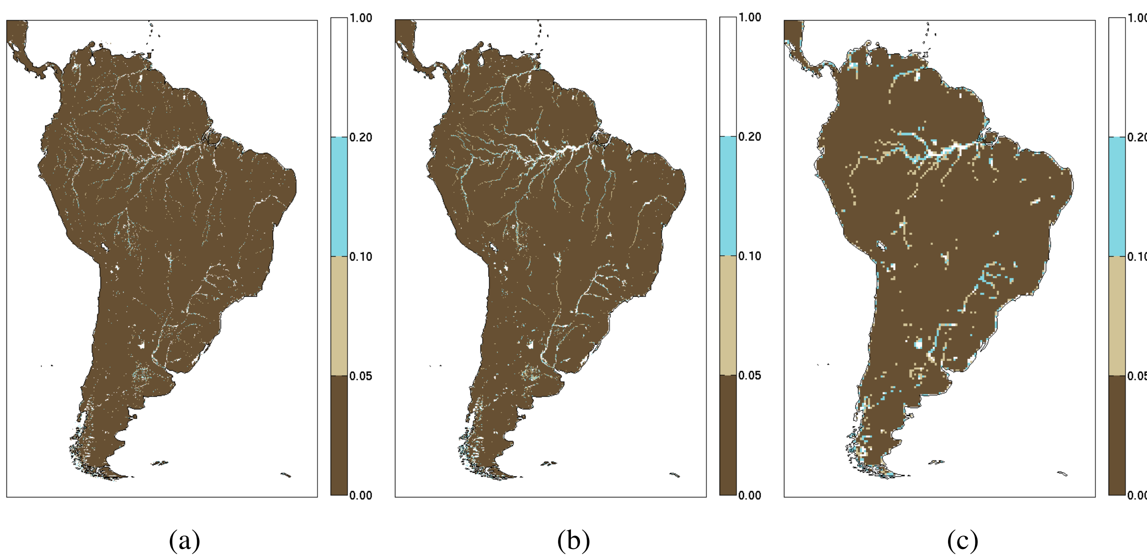

Figure 4. Static water fraction at different spatial scales over South America at (a) 3 km, (b) 9 km, and (c) 36 km grid resolutions on the EASE-Grid 2.0 projection.

Figure 4 shows a close-up over South America illustrating details of the distinction between the global static water fraction maps at the different grid resolution scales. Water fraction varies with the spatial scale. For example, globally there are fewer 36-km grid cells containing 50% or more water than 3-km grid cells of the same water fraction. Figures 5-7 are combined PDF/CDF plots of the number of land grid cells[1] versus water fraction at 3, 9, and 36 km grid resolutions. It is evident that, on a global basis, **the majority of land pixels (over 80%) contain less than 10% water regardless of spatial scale.**

[1] A land grid cell is defined as one that does not contain 100% water.

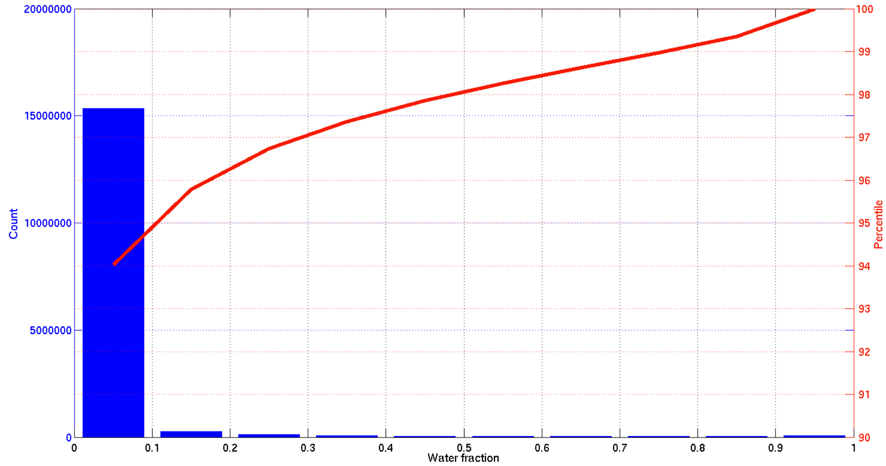

Figure 5. Pareto plot of number of 3-km land grid cells versus water fraction

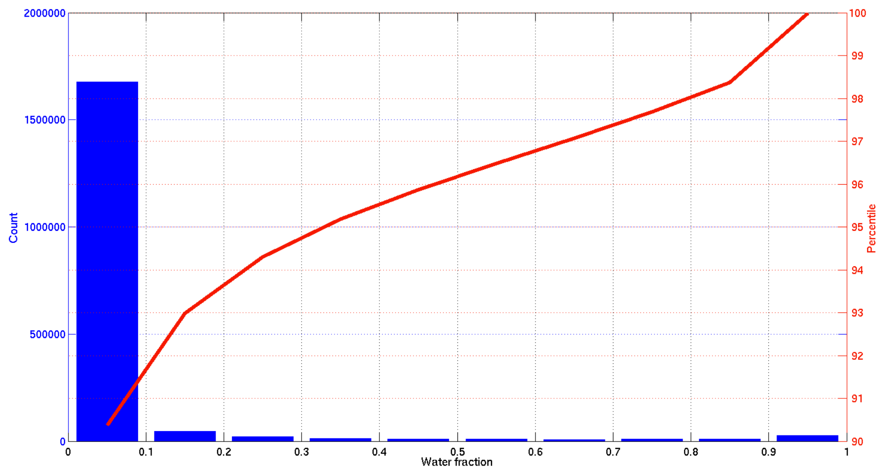

Figure 6. Pareto plot of number of 9-km land grid cells versus water fraction.

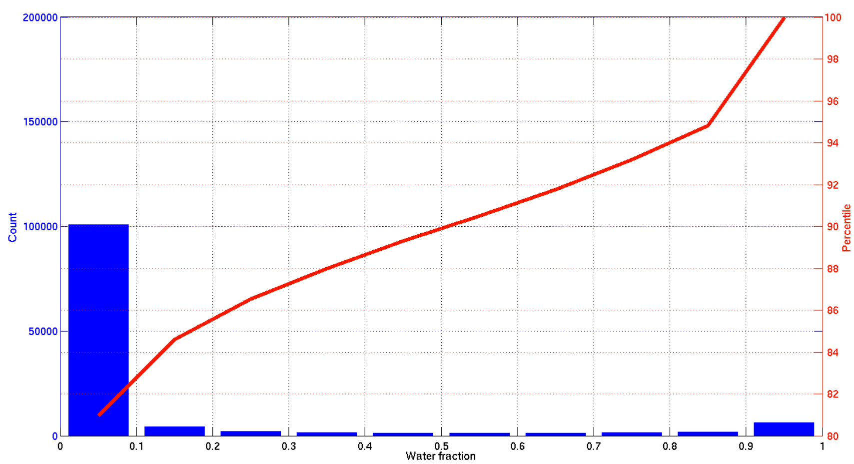

Figure 7. Pareto plot of number of 36-km land grid cells versus water fraction.

3.4 Dataset Specifications

The dataset specifications for the 1, 3, 9, and 36 km EASE-grid output data files are provided in Appendix C.

4 Conclusion

The search for an appropriate source dataset to use in deriving water fraction data for SMAP was aided by work conducted by the MODIS Land Science Team. This team had realized the same need for their land products and had developed a high-resolution water/land mask product in 2008. The result was the MOD44W 250-m water/land mask, which addressed many shortcomings of earlier products (e.g., water discontinuity, outdated coastline database, etc.). Because of the many improvements realized by MOD44W the MODIS Land Science Team recommended its use as the standard water mask for the MODIS Collection 6 reprocessing (personal communication, Mark Carroll, lead developer of the MOD44W dataset). We therefore are confident in adopting MOD44W as the source dataset to derive the SMAP water fraction datasets.

5 Acknowledgment

This work was carried out at the Jet Propulsion Laboratory, California Institute of Technology, under contract with the National Aeronautics and Space Administration.

6 References

Carroll, M., J. Townshend, C. DiMiceli, P. Noojipady, and R. Sohlberg (2009): A New Global Raster Water Mask at 250-Meter Resolution, *International Journal of Digital Earth, 2(4),* 291-308.

Haran, T., J. Bohlander, T. Scambos, T. Painter, and M. Fahnestock (2006): *MODIS mosaic of Antarctica (MOA) image map*. Boulder, Colorado USA: National Snow and Ice Data Center. Digital media, http://nsidc.org/data/nsidc-0280.html. (Accessed Mar 6, 2012.)

Homer, C., C. Huang, L. Yang, B. Wylie and M. Coan (2004): Development of a 2001 National Land Cover Database for the United States. *Photogrammetric Engineering and Remote Sensing*, *70(7),* 829-840., http://www.mrlc.gov/nlcd2001.php. (Accessed Mar 6, 2012.)

NASA (2007): MODIS 16-day Composite MOD44C, latlon.na.2004289b3, Collection 4, The Global Land Cover Facility, University of Maryland, College Park, Maryland, Day 289, 2004.

REVERB (2012): REVERB: The Next Generation Earth Discovery Tool, http://reverb.echo.nasa.gov/reverb/. (Accessed Mar 6, 2012.)

Rodriguez, E., C. S. Morris, J. E. Belz (2006): A global assessment of the SRTM performance, *Photogrammetry Engineering and Remote Sensing*, *72,* 249-260.

SWBD (2012): The Shuttle Radar Topography Mission Water Body Data Set, http://www2.jpl.nasa.gov/srtm/index.html. (Accessed Mar 6, 2012.)

Appendix A: SMAP Science Data Products and ATBDs

The SMAP Algorithm Theoretical Basis Documents are available at the SMAP web site http://smap.jpl.nasa.gov/science/dataproducts/ATBD/.

Data Product	Description	ATBD
L1A_Radar	Radar raw data in time order	(Joint with L1C_S0_HiRes)
L1A_Radiometer	Radiometer raw data in time order	(Joint with L1B_TB)
L1B_S0_LoRes	Low resolution radar σ_o in time order	(Joint with L1C_S0_HiRes)
L1C_S0_HiRes	High resolution radar σ_o (half orbit, gridded)	West, R., L1B & L1C radar products, JPL D-53052, JPL, Pasadena, CA.
L1B_TB	Radiometer T_B in time order	Piepmeier, J. et al., L1B radiometer product, GSFC SMAP-006, GSFC, Greenbelt, MD.
L1C_TB	Radiometer T_B (half orbit, gridded)	Chan, S. et al., L1C radiometer product, JPL D-53053, JPL, Pasadena, CA.
L2_SM_A	Soil moisture (radar, half orbit)	Kim, S. et al., L2 & L3 radar soil moisture (active) product, JPL D-66479, JPL, Pasadena, CA.
L2_SM_P	Soil moisture (radiometer, half orbit)	O'Neill, P. et al., L2 & L3 radiometer soil moisture (passive) product, JPL D-66480, JPL, Pasadena, CA.
L2_SM_AP	Soil moisture (radar/radiometer, half orbit)	Entekhabi, D. et al., L2 & L3 radar/radiometer soil moisture (active/passive) products, JPL D-66481, JPL, Pasadena, CA.
L3_FT_A	Freeze/thaw state (radar, daily composite)	McDonald, K. et al., L3 radar freeze/thaw (active) product, JPL D-66482, JPL, Pasadena, CA.
L3_SM_A	Soil moisture (radar, daily composite)	(Joint with L2_SM_A)
L3_SM_P	Soil moisture (radiometer, daily composite)	(Joint with L2_SM_P)
L3_SM_AP	Soil moisture (radar/radiometer, daily composite)	(Joint with L2_SM_AP)
L4_SM	Soil moisture (surface & root zone)	Reichle, R. et al., L4 surface and root-zone soil moisture product, JPL D-66483, JPL, Pasadena, CA.
L4_C	Carbon net ecosystem exchange (NEE)	Kimball, J. et al., L4 carbon product, JPL D-66484, JPL, Pasadena, CA.

Appendix B: SMAP Ancillary Data Reports

The SMAP Ancillary Data Reports are available with the ATBDs at the SMAP web site http://smap.jpl.nasa.gov/science/dataproducts/ATBD/.

Data/Parameter	Ancillary Data Report
Crop Type	Kim, S., Crop Type, JPL D-53054, Pasadena, CA
Digital Elevation Model	Podest, E. et al., Digital Elevation Model, JPL D-53056, Pasadena, CA
Landcover Classification	Kim, S., Landcover Classification, JPL D-53057, Pasadena, CA
Soil Attributes	Das, N. et al., Soil Attributes, JPL D-53058, Pasadena, CA
Static Water Fraction	Chan, S. et al., Static Water Fraction, JPL D-53059, Pasadena, CA
Urban Area	Das, N., Urban Area, JPL D-53060, Pasadena, CA
Vegetation Water Content	Chan, S. et al., Vegetation Water Content, JPL D-53061, Pasadena, CA
Permanent Ice	McDonald, K., Permanent Ice & Snow, JPL D-53062, Pasadena, CA
Precipitation	Dunbar, S., Precipitation, JPL D-53063, Pasadena, CA
Snow	Kim, E. et al., Snow, GSFC SMAP-007, Greenbelt, MD
Surface Temperature	Fisher, J. et al., Surface Temperature, JPL D-53064 Pasadena, CA
Vegetation and Roughness Parameters	Colliander, A., Vegetation & Roughness Parameters, JPL D-53065, Pasadena, CA

Appendix C: Dataset Specifications

Based on the processing steps outlined in Section 3, the MOD44W water/land mask data product is converted to global maps of static water fraction at 1, 3, 9, and 36 km grid resolutions on the EASE-Grid 2.0 projection.

1-km global static water fraction database	
Filename	waterfrac01km.14616x34704.float32
Location	/project/ancillary/versions/landwater/mod44w
Precision	4-byte floating point in little-endian byte order
Arrangement	14616 rows by 34704 columns written in column-major order
Data unit	Dimensionless
Data minimum	0.0 (no water)
Data maximum	1.0 (all water)
Map projection	EASE-Grid 2.0

3-km global static water fraction database	
Filename	waterfrac03km. 4872x11568.float32
Location	/project/ancillary/versions/landwater/mod44w
Precision	4-byte floating point in little-endian byte order
Arrangement	4872 rows by 11568 columns written in column-major order
Data unit	Dimensionless
Data minimum	0.0 (no water)
Data maximum	1.0 (all water)
Map projection	EASE-Grid 2.0

9-km global static water fraction database	
Filename	waterfrac09km.1624x3856.float32
Location	/project/ancillary/versions/landwater/mod44w
Precision	4-byte floating point in little-endian byte order
Arrangement	1624 rows by 3856 columns written in column-major order
Data unit	Dimensionless
Data minimum	0.0 (no water)
Data maximum	1.0 (all water)
Map projection	EASE-Grid 2.0

36-km global static water fraction database	
Filename	waterfrac36km.406x964.float32
Location	/project/ancillary/versions/landwater/mod44w
Precision	4-byte floating point in little-endian byte order
Arrangement	406 rows by 964 columns written in column-major order
Data unit	Dimensionless
Data minimum	0.0 (no water)
Data maximum	1.0 (all water)
Map projection	EASE-Grid 2.0